Ca' Foscari
University
of Venice

**Department
of Economics**

Working Paper

Alessandro Moro

**Distribution Dynamics of
Property Crime Rates in
the United States**

ISSN: 1827-3580
No. 13/WP/2016

Distribution Dynamics of Property Crime Rates in the United States

Alessandro Moro

Ca' Foscari University of Venice

Abstract

Using crime data for the 48 continental and conterminous US states and the distribution dynamics approach, this paper detects two distinct phases in the evolution of the property crime distribution: a period of strong convergence (1971-1980) is followed by a tendency towards divergence and bimodality (1981-2010). Moreover, the analysis reveals that differences in income per capita and police can explain the emergence of a bimodal shape in the distribution of property crime: in fact, after conditioning on these variables, the bimodality completely disappears. This empirical evidence is consistent with the predictions of a two-region model, that stresses the importance of income inequality in determining the dynamics of the property crime distribution.

Keywords

Crime Distribution, Property Crime, Convergence Analysis, Distribution Dynamics, Non-Parametric Statistics

JEL Codes

C14, C63, C72

Address for correspondence:

Alessandro Moro

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
Fax: (+39) 041 2349176
e-mail: alessandro.moro@unive.it

This Working Paper is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.

Distribution Dynamics of Property Crime Rates in the United States

Alessandro Moro

Department of Economics,

Ca' Foscari University, Cannaregio 873, 30121 Venice.

alessandro.moro@unive.it.

Abstract

Using crime data for the 48 continental and conterminous US states and the distribution dynamics approach, this paper detects two distinct phases in the evolution of the property crime distribution: a period of strong convergence (1971-1980) is followed by a tendency towards divergence and bimodality (1981-2010). Moreover, the analysis reveals that differences in income per capita and police can explain the emergence of a bimodal shape in the distribution of property crime: in fact, after conditioning on these variables, the bimodality completely disappears. This empirical evidence is consistent with the predictions of a two-region model, that stresses the importance of income inequality in determining the dynamics of the property crime distribution.

Keywords

Crime Distribution, Property Crime, Convergence Analysis, Distribution Dynamics, Non-Parametric Statistics

JEL Codes: C14, C63, C72

1. Introduction

A huge literature has tried to understand the causes and the dynamics of crime and important contributions can be found in different social science disciplines: from economics (Becker, 1968) to sociology (Cohen and Felson, 1979), from politics (Smith, 1997) to law (Marvell and Moody, 2001). A huge set of determinants has been considered to explain crime, including unemployment, income inequality, police officers, the legalization of abortion, and many others: see, *inter alia*, Cornwell and Trumbull (1994), Marvell and Moody (1996), Kapuscinski *et al.* (1998), Donohue and Levitt (2001), Greenberg (2001), Paternoster and Bushway (2001), Raphael and Winter-Ebmer (2001), Gould *et al.* (2002), Levitt (2002), Machin and Meghir (2004), Baltagi (2006), Rosenfeld and Fornago (2007).

Moreover, the empirical crime literature has focused its attention on the dynamics of the aggregate level of crime in the US, such as Cook and Cook (2011). In particular, many papers have tried to explain the raise of crime rates in the 1970s and 1980s and their decline in the 1990s. Reviewing the evidence from these numerous papers, Levitt (2004) identifies four factors that explain the decrease of crime started in the 1990s: the increasing number of police, the skyrocketing number of prisoners, the ebbing of the crack epidemic and the legalization of abortion in the 1970s. Another paper trying to explain the dynamics of crime is represented by Shoesmith (2010), who uses an error correction model and four factors to explain both the raise and fall of crime rates. The employed factors are arrest rates, income per capita, the proportion of justice resources devoted to drug crime, and alcohol consumption.

The aim of this paper is different because it does not focus on the aggregate value of crime but on the analysis of the entire distribution of crime rates for the US states and on its evolution over time. Other papers have tried to address this issue using standard approaches, like the beta-convergence and sigma-convergence concepts: for instance, Cook and Winfield (2013), employing growth regressions and measures of cross-sectional variation, provide evidence of convergence for the US states, whereas, using

a different probabilistic approach, Cook and Watson (2013) find different results depending on the period analysed and on the crime typology considered.

However, it has been recognised that a negative relationship between initial values and growth rates, which is the essence of the beta-convergence analysis (Baumol, 1986; Barro, 1991; Barro and Sala-i-Martin, 1991, 1992), is a necessary but not a sufficient condition for convergence (Quah, 1993a): in fact, the regression approach, by focusing on the behaviour of a representative unit towards its own steady state, is completely silent on what happens to the entire cross-sectional distribution. Moreover, the sigma-convergence analysis (Sala-i-Martin, 1996) is also affected by significant drawbacks because, as argued by Quah (1996a), a constant variability is compatible with very different dynamics, from criss-crossing and leap-frogging to persistent inequality. Distinguishing between these completely different patterns is crucial and is possible only by analysing the entire cross-sectional distribution.

In order to overcome these criticisms affecting the existing literature about convergence analysis of crime, the present paper adopts an alternative methodology, i.e. the distribution dynamics approach (Quah, 1993a and b, 1996a and b, 1997), which allows the study of the entire distribution of crime rates, both in terms of shape and of intra-distributional dynamics.

In particular, this paper considers the dynamics of property crime rates for 48 conterminous US states during the period 1971-2010. Results indicate different phases for this typology of crime: in fact, a period of strong convergence (1971-1980) is followed by a tendency towards divergence and bimodality (1981-2010).

Furthermore, the analysis reveals that differences in the levels of income per capita, whose divergence starting from the 1980s has been documented by Gerolimetto and Magrini (2014) in a distribution dynamics setting, and in the numbers of state police employees, which also exhibit a divergent pattern

starting from the 1980s, can explain the emergence of a bimodal shape in the distribution of property crime rates: in fact, after conditioning on these two variables, the bimodality completely disappears.

The rest of the paper is organised as follows: Section 2 presents a two-region model which investigates the relationships between crime and other socio-economic variables from a theoretical point of view; Section 3 introduces the adopted empirical methodology, whereas Section 4 describes the data employed; Sections 5-7 present the empirical results obtained with the distribution dynamics technique and explore the relationship between crime and some of its determinants; finally, Section 8 concludes.

2. A Two-Region Model of Crime Dynamics

This section presents a theoretical framework in which it is possible to explore the dynamics of crime over time and across different spatial units.

In particular, in this model there are two regions or states, indicated by A and B . In each state $i \in \{A, B\}$ and at each time $t \in \{1, 2, \dots\}$, the number of property crimes committed is denoted by $c_{i,t}$. The evolution of property crime in each region depends on its past value $c_{i,t-1}$, on the number of police resources $p_{i,t}$ and on the level of after-tax income $(1 - \tau_{i,t})y_{i,t}$:

$$c_{i,t} = (1 + \rho)c_{i,t-1} - \psi p_{i,t} - \sigma(1 - \tau_{i,t})y_{i,t} \quad (1)$$

where ρ is the growth rate of crime when police and income are both zero, while ψ and σ measure how much crime is sensible to current variations of police resources and after-tax income, respectively. The reason for this specification is very intuitive: a higher level of police resources means more repression activities, whereas a higher value of legal income makes crime less attractive. Moreover, this specification is consistent with economic models of crime, such as Becker (1968), which show that

individuals tend to commit crime when the probability of punishment or the opportunities in the legal labour market are low.

On the other hand, the future value of income depends on the current level of crime, because the presence of high crime rates reduces the resources available in each state by discouraging investments:

$$y_{i,t+1} = (1 + \theta)y_{i,t} - \omega c_{i,t} + \eta_{i,t+1} \quad (2)$$

in which θ is the growth rate of income in the absence of crime, ω measures how much income is sensible to current variations of crime, and, finally, $\eta_{i,t+1}$ is an exogenous, independently distributed shock with zero-mean, capturing all the future positive or negative surprises (e.g. unexpected increase or decrease of labour productivity).

The two states decide, through taxation, the share of resources $\tau_{i,t}$ to allocate to police:

$$p_{i,t} = \tau_{i,t} y_{i,t} \quad (3)$$

Consequently, the expenditure in crime repression improves the future level of income by reducing crime and its negative effect on the level of resources, but it is a costly investment because it requires an increase of the tax rate.

It is assumed that politicians choose in each period the level of taxation by maximising the probability of being re-elected Π , which is a strictly increasing function of the expected economic performance minus the current costs of taxation (for simplicity, elections take place in each period):

$$\max_{\tau_{i,t}} \Pi \left(\delta E_{i,t}(y_{i,t+1}) - \phi \frac{\tau_{i,t}^2}{2} \right) \quad (4)$$

where $\Pi'(\cdot) > 0$, $E_{i,t}(\cdot)$ is the expected value conditioned on the information set available at time t in state i , while δ and ϕ are the two weights associated to the expected future economic performance $E_{i,t}(y_{i,t+1})$ and the political costs of taxation $\frac{\tau_{i,t}^2}{2}$, respectively.

The first order conditions of the problem can be rewritten in the following way:

$$\phi \tau_{i,t} = \omega \delta (\psi - \sigma) y_{i,t} \quad (5)$$

in which the left-hand side represents the marginal cost of taxation, whereas the right-hand side is the marginal benefit. It is worth noting that taxation has a positive benefit only if the crime sensitivity to police, measured by ψ , is higher than the degree of dependence on after-tax income, indicated by σ .

The optimal level of taxation equalizes marginal benefits and marginal costs:

$$\tau_{i,t}^* = \frac{\omega \delta (\psi - \sigma) y_{i,t}}{\phi} \quad (6)$$

This expression shows that the optimal tax is higher when: the weight associated to taxation, measured by ϕ , is lower; the net benefit of taxation $(\psi - \sigma) y_{i,t}$ is higher; income is more sensitive to crime, as indicated by ω ; the effect of the economic performance on the probability of being re-elected δ is higher. Substituting the last quantity in the equation of crime, income and police, it is possible to obtain the optimal dynamics for these three variables. In the model calibration, it is reasonably assumed that the net benefit of taxation is positive ($\psi = 0.08, \sigma = 0.05$) and that the probability of being re-elected is more sensible to the economic performance than the cost of taxation ($\delta = 0.95, \phi = 0.35$).

This model allows to analyse the effects of an exogenous income shock and its propagation to the other variables of the system. In particular, the case of an increase in income inequality is considered because this scenario is related to the increasing inequality in terms of income per capita observed across the US

states, starting from the 1980s. Consequently, it is assumed in the main calibration that state A is the richest region ($y_{A,0} = 18 > y_{B,0} = 16$), which also starts with the highest level of crime ($c_{A,0} = 13 > c_{B,0} = 11$). The growth rates of income and crime are $\theta = 0.050$ and $\rho = 0.085$, respectively, and the sensitivity of income to crime is $\omega = 0.06$. The two regions are followed for 40 periods to mimic the forty years considered in the next empirical analysis, i.e. the period from 1971 to 2010. It is assumed that the richest region has the highest number of crimes committed because this is consistent with the data, showing a positive correlation between property crime and income per capita at the beginning of the 1970s. Moreover, the shocks $\eta_{i,t}$ are always zero, except for region A at time 10, and the analysis considers a low (1.8), medium (2.6) and high value (3.8) of the shock $\eta_{A,10}$.

Figure 1 reports the dynamics implied by the model for the main variables in the different scenarios considered. In particular, the columns correspond to the small, medium and big shock scenario, respectively. The first row shows the timing and magnitude of the shock; the second and third rows present the dynamics over time of income and crime in the two regions, compared to the average of the two states in each period, respectively (the time series of tax rate and police have the same pattern of income and, for this reason, are omitted); finally, the last row depicts the trajectory of the aggregate value of crime, which is the sum of the crime levels in the two regions. In every scenario the mechanism behind the model is the same: a higher (lower) level of income leads to a decline (raise) of crime through both the opportunity cost and the police channel.

In the case of a small shock affecting the income of the richest region, this state has not enough resources to reduce its crime level, which continues to grow and subtract wealth to the state, and this leads to a divergent crime pattern. Moreover, the aggregate level of crime exhibits an explosive trend due to the strong growth of crime in both regions.

On the other hand, the second column shows that a bigger income shock provides to state A enough resources to limit the growth of its crime level: as a consequence, the time series of crime in the two regions converge to their average at the end of the forty periods considered. However, the aggregate level of crime presents an increasing linear trend.

Finally, with an even bigger shock, the dynamics of the system changes completely: in this scenario the richest region is able to reduce its crime level compared to the poorest region. Benefiting from this crime reduction, income in region A displays an increasing pattern, which also leads to a greater inequality in the available resources. Since the richest region lowers its crime level more than the poorest state, the result is that the two crime trajectories converge in a first phase, but then they diverge in the final periods. Furthermore, looking at the aggregate value of crime in the two states, an increasing trend in the first periods is followed by a decline, because the decrease of crime in region A is larger than the increase in state B .

If the values assigned to the weight of taxation ϕ , the income sensitivity to crime ω , or the crime growth rate ρ are increased, results gradually change towards a divergence of the distribution of crime in the two states and to an explosive trend in the aggregate level of crime. The same happens if the crime sensitivity to police ψ and after-tax income σ , the effect of economic performance on the probability of re-election δ , or the income growth rate θ are reduced.

The three scenarios considered reveal that, from a theoretical point of view, a greater income inequality, such as the one experienced in the US starting from the 1980s, can have completely different consequences on the dynamics of crime and on its distribution, depending on the size of the shocks that drive this phenomenon. The following empirical analysis will reveal which theoretical scenario best approximates the US experience.

3. Methodology

As mentioned in the introductory section, there are two approaches to the analysis of convergence: the regression method, complemented by the study of the cross-sectional variability, and the distribution dynamics approach. In this paper the latter approach is chosen because it allows the study of the entire cross-sectional distribution of a given variable, both in terms of external shape and intra-distributional dynamics, using stochastic kernels to describe its evolution over time.

Let c_t and c_{t+s} represent the crime rates, relative to the group average, of a set of n states at time t and $t + s$, respectively. Moreover, assuming that the two random variables admit a density, denoted by $f(c_t)$ and $f(c_{t+s})$, and that $f(\cdot)$ can be modeled as a first order process, the density at time $t + s$ is given by:

$$f(c_{t+s}) = \int_{-\infty}^{+\infty} f(c_{t+s}|c_t)f(c_t)dc_t \quad (7)$$

in which $f(c_{t+s}|c_t)$ is the conditional density function that maps the values at time t into the values at time $t + s$. This density function is called stochastic kernel and it plays a crucial role within this approach: in fact, it provides information on the movements from one part of the distribution to another one between time t and $t + s$. For this reason, once the stochastic kernel is estimated, convergence can be analysed directly from its shape or, assuming a time homogeneous markov process for the studied phenomenon, by comparing the initial distribution at time t to the so-called ergodic distribution, which is the limit of $f(c_{t+s})$ when s goes to infinity. In more intuitive terms, the ergodic distribution is the characterisation of the likely long-run cross-sectional distribution of the variable of interest.

The output of this type of analysis is represented by a set of figures: (i) a three-dimensional plot of the estimated stochastic kernel (the Technical Appendix sketches the estimation procedure employed) ; (ii) the corresponding contour plot, with contours at the 90%, 50% and 10% level; (iii) the Highest Density

Regions (HDR) plot, proposed by Hyndman (1996), in which the vertical strips represent conditional densities given specific values for the initial year and, for each strip, darker to lighter areas display the 10%, 50% and 90% highest density regions; (iv) a plot comparing the initial year distribution with the final year one and the ergodic.

Then, convergence is analysed by looking at the three-dimensional shape of the stochastic kernel and at the corresponding contour and HDR plots or by comparing the initial distribution with the final one and the ergodic. A probability mass located along the main diagonal in the contour and HDR plots indicates a persistence feature of the studied phenomenon, because the elements in the cross-sectional distribution remain where they started. On the other hand, a convergence process is highlighted by a probability mass concentrated around the mean value at time $t + s$ and parallel to the time t axis, while a distribution parallel to the $t + s$ axis is a signal of divergence: in fact, in the first case, the units of analysis, characterised by different values at time t , will exhibit a similar value at time $t + s$, whereas, in the second case, the opposite behaviour is observed. Finally, the formation of two, or more, modes in the ergodic distribution reveals a tendency towards polarization or stratification.

4. Data

The crime data used in this paper are from the Federal Bureau of Investigation (FBI)'s Uniform Crime Reports (UCR)¹, which include crime reports submitted voluntarily either directly by local, state, federal or tribal law enforcement agencies or through centralised state agencies across the country. Data are freely available online for each year, starting from 1960, and for the 50 US states, plus the District of Columbia. Crimes are classified according to the following categories: burglary, larceny-theft, motor-vehicle theft, murder, rape, robbery, assault. The first three typologies are grouped in the property crime category while the others are classified as violent crimes. The attention of this paper is devoted to the

aggregate category of property crime, which is measured by a standard index, the property crime rate, defined as the number of reported property crimes committed per 100,000 inhabitants.

Figure 2 plots the time series of the property crime rate in the US. This picture shows a very well-known pattern in the crime literature: the aggregate level of property crime increases up to the beginning of the 1990s and then falls. It is worth stressing that this trend seems consistent with the predictions of the model in the third theoretical scenario of Figure 1.

The UCR data have many advantages: they cover a long period of time with a stable methodology, allowing a meaningful trend analysis; they are the only source of geographically disaggregated crime data available for the US; they offer a good coverage in terms of crime typologies and in terms of geographic locations considered. On the other hand, the UCR program has some limitations: it covers only crime reported to the police and many crimes are reported in low percentages; furthermore, since reporting is voluntary, some enforcement agencies may not report information or information may be incomplete. In the presence of incomplete data, the FBI uses specific protocols to impute the missing values: the imputation is based on crime rates of agencies considered similar according to population size, type of agencies (city, rural and state, suburban counties) and geographic location.²

In order to bring the model to the data, a measure of income and police is needed. The relationship between property crime and income will be explored in the following analysis using data on real per capita personal income: the personal per capita income net of current transfer receipts comes from the Bureau of Economic Analysis,³ while the consumer price index (CPI) used to deflate income is from the Bureau of Labor Statistics.⁴ On the other hand, the level of crime repression activities in each state is measured by the number of state police employees per 100,000 inhabitants, as recorded in the UCR data.⁵ Furthermore, our analysis is restricted to the forty-year period from 1971 to 2010⁶ and to the 48 continental and conterminous US states.⁷

5. Distribution Dynamics of Property Crime

Using the methodology described in Section 3, the dynamics over time of the distribution of property crime is analysed.

Figure 3 shows the results for the overall period, from 1971 to 2010. The first three graphs are related to the shape of the estimated stochastic kernel. In particular, if we look at the contour levels and at the 45 degree line, it is possible to analyse the intra-distributional dynamics between the two considered periods: observations with a high property crime rate in 1971 are likely to have a lower crime rate in 2010; states with a low rate in the initial year tend to present a higher crime rate in the final year; moreover, there is a group of states located around the mean in the initial year experiencing the highest crime rates in the final period. This situation results in a final and ergodic distribution affected by bimodality, as displayed in the fourth graph.⁸ The states that are responsible for the first mode of the final density are those located in the north-west part of the US while the second mode can be attributed to the states in the south-east.

However, behind this dynamics, two distinct phases are identifiable. In fact, Figure 4 depicts a period of strong convergence from 1971 to 1980, made evident by a sharply concentrated unimodal ergodic distribution and by a noticeable clockwise rotation of the estimated probability mass: this rotation represents evidence of convergence because it implies that states with a low crime rate at the beginning of the period considered present higher rates at the end; and the opposite holds for states with high crime rates. Conversely, Figure 5 shows a phase of divergence, from 1981 to 2010, in which the distribution of property crime rates shows a clear tendency towards bimodality, made transparent by the comparison of the initial distribution with the final one and the ergodic.

6. Robustness Checks

These findings seem to suggest a dynamics more similar to the third scenario predicted by the model. In order to test the robustness of the obtained results and to address the concern that a particular city or county is driving the whole of the distribution (e.g. the city of New York in the state of New York), the analysis is replicated considering the 9 regions in which the US territory is divided according to the definition of the United States Census Bureau.

Figure 6 presents the distribution dynamics analysis for these regions in the period 1971-1980: convergence is highlighted by a clockwise rotation of the probability mass and by the transition from an initial distribution characterised by high variability to a more concentrated, although bimodal, final distribution. A further evidence of convergence is represented by the ergodic distribution which exhibits lower variability than the final distribution and in which any evidence of bimodality disappears. Figure 7 replicates this exercise for the period 1981-2010: the last picture in this figure reveals the emergence of a final and ergodic distribution far less concentrated than the initial one. With only 9 regions it is not possible to capture the bimodal shape observed in Figure 5; however, the tendency towards divergence is clear.

The analysis performed on more aggregated regional divisions proves the robustness of the results regarding the existence of two phases in the evolution of the distribution of property crime, i.e. a period of convergence followed by a divergent dynamics. This evidence is again consistent with the third scenario predicted by the model, characterised by the presence of strong income shocks.

7. Conditional Distribution Dynamics

According to the theoretical framework of Section 2, the divergence of property crime rates is driven by income shocks that affect crime through the channels of opportunity cost and police. Moreover, in the same period in which a tendency towards bimodality has been detected for property crime rates, i.e. the years from 1981 to 2010, Gerolimetto and Magrini (2014) find an analogous divergent pattern using real per capita personal income for the same 48 conterminous US states: in fact, the ergodic distribution in the last picture of Figure 8 presents a higher concentration in the right tail than the initial one. By applying the distribution dynamics technique, the same divergent tendency can be observed in the number of state police employees, starting from the 1980s: Figure 9 shows this feature. These considerations motivate the idea to explore more deeply the relationship between income, police and crime from an empirical point of view.

In order to understand if income and police can explain the bimodality of property crime, the conditional distribution of the variable of interest must be considered. Since both income and police are likely to be endogenous variables,⁹ the method proposed by Quah (1996b) is considered in the following analysis.

The first step of the procedure consists in estimating a growth regression, in which the growth rate of the dependent variable (property crime) is regressed on a set of potentially endogenous variables (income and police), and then taking the fitted values for subsequent analysis: in order to correct for the possible endogenous nature of these variables and, in particular, the presence of feedback effects, both the current values, the lags and the leads are included in the regression. The inclusion of leads might not solve completely the endogeneity problem due to the many possible channels of reversed causality between dependent and independent variables: this means that the following results shall be interpreted with caution and on a descriptive basis.

In the second step of the procedure, the fitted values are used to estimate the residual component of crime that is not explained by the accumulation of the conditioning variables (details are in the Technical Appendix). Figure 10 and 11 are obtained by applying the distribution dynamics analysis to this unexplained component of property crime, for the period 1971-2010 and 1981-2010, respectively.

The conditional convergence is made evident in both figures by the noticeable clockwise rotation of the estimated probability mass. Moreover, any sign of bimodality is completely disappeared from the ergodic distribution, which is also more concentrated than the initial and final one. Thus, the divergence of the number of state police forces and the increasing income inequality, started in the 1980s, explain well the tendency towards a bimodal shape of the property crime distribution.

8. Concluding Remarks

The importance of analysing the convergence of crime rates, as stated by Cook and Winfield (2013), is related to the existence of a possible national crime trend and whether there are movements towards it. Moreover, the analysis of convergence helps in choosing between competing sociological theories, like the modernization and the conflict theories (LaFree, 2005): in fact, according to the modernization view, crime rates should converge given the spread of developments and advances across regions, whereas the conflict theory predicts their divergence, arguing that these developments have uneven speed in the different regions.

The proposed theoretical model predicts the emergence of two distinct phases in the dynamics of property crime distribution when spatial units are affected by strong income shocks generating greater inequality: a period of convergence followed by the divergence of crime rates. This scenario fits the situation of the US states, because economic disparities were exacerbated in those regions, starting from the 1980s.

This theoretical prediction is confirmed by the descriptive analysis presented in this paper. In fact, using the distribution dynamics methodology, two different patterns are identified during the period 1971-2010 for the property crime distribution: a phase of strong convergence (1971-1980), followed by a period of divergence and a tendency towards bimodality (1981-2010). These two distinct phases, to the best of our knowledge, have not been highlighted by the existing literature.

Moreover, the contemporaneous divergence of personal per capita income and of state police rates can account for the observed dynamics of property crime rates: in fact, conditioning on income per capita and state police, the distribution of property crime does not exhibit a bimodal shape, thus indicating the presence of conditional convergence.

An important policy implication that can be derived from the analysis is as follows. The theoretical framework suggests that significant income disparities are translated into different concentrations of crime through the opportunity cost and the police channel. Moreover, the model predicts that poor states have lower resources to fight crime and, consequently, they exhibit higher crime rates. Since the presence of crime discourages investments and lowers income, these states are trapped in a vicious circle. Therefore, mitigating the effects of inequality with cross-state compensations, in terms of financial and police resources, may help avoiding both the concentration of crime activities in specific regions and the emergence of self-reinforcing gaps between poor and rich states.

Technical Appendix

A1. Stochastic Kernel Estimation

Observing a given variable c on a set of n statistical units at two points in time, t and $t + s$, the easiest way to estimate the stochastic kernel is through the kernel density estimator, which takes the following expression:

$$\hat{f}(c_{t+s}|c_t) = \frac{\hat{f}(c_{t+s}, c_t)}{\hat{f}(c_t)} \quad (A1.1)$$

where the kernel estimator of the joint density function is:

$$\hat{f}(c_{t+s}, c_t) = \sum_{i=1}^n \frac{1}{nh_{c_t}h_{c_{t+s}}} K\left(\frac{c_t - c_{i,t}}{h_{c_t}}\right) K\left(\frac{c_{t+s} - c_{i,t+s}}{h_{c_{t+s}}}\right) \quad (A1.2)$$

and the kernel density estimator of the initial density is:

$$\hat{f}(c_t) = \sum_{i=1}^n \frac{1}{nh_{c_t}} K\left(\frac{c_t - c_{i,t}}{h_{c_t}}\right) \quad (A1.3)$$

in which $K(\cdot)$ is a kernel function and h is a chosen bandwidth parameter that controls the degree of smoothing applied to the density estimate.

However, as argued by Hyndman *et al.* (1996), this estimator might have poor properties in terms of bias: the authors propose to adjust the estimate of the mean function implicit in the kernel density estimator with one obtained from a smoother with better bias properties. For this reason, according to Loader (1999), the local linear estimator is chosen in the present analysis. Following Gerolimetto and Magrini (2014), in order to estimate the conditional density, a Gaussian kernel is selected with a nearest-neighbour bandwidth in the initial year dimension, with a span of 30% of the data, and a fixed Normal Scale

bandwidth, as suggested by Silverman (1986), in the final year dimension. The local linear estimator for the mean function has a nearest-neighbour bandwidth with a span chosen to minimize the AIC criterion.

A2. Conditioning Method

The conditioning analysis performed adopts and adapts the method proposed by Quah (1996b). In the first step of the procedure, the growth rate of the property crime rate is regressed on the growth rates of personal per capita income and state police rates, current, lagged and leaded:

$$g_{i,t}^c = \alpha + \sum_{j=-l}^l (\beta_j g_{i,t+j}^y + \gamma_j g_{i,t+j}^p) + \varepsilon_{i,t} \quad (A2.1)$$

where $g_{i,t}^x$ is the growth rate of the variable x between time $t - 1$ and t in state i , whereas l represents the number of lags and leads considered in the regression (in the present analysis l is equal to three).

The second step of the procedure consists in the accumulation, state by state, of the fitted values taken from the previous regression, to obtain the time-varying trend paths explained by the conditioning variables:

$$g_{i,t}^* = \sum_{j=1}^t \hat{g}_{i,j}^c \quad (A2.2)$$

This determines up to a multiplicative time-invariant level the component for each state explained by income and police: as in Quah (1996b), it is assumed that this state-specific multiplicative level is equal to a linear combination of the time averages of income $\bar{y}_i$ and police $\bar{p}_i$. The parameters of this linear combination are determined by solving the following minimisation problem:

$$\min_{a,b} \sum_i \sum_t [c_{i,t} - (a\bar{y}_i + b\bar{p}_i)(1 + g_{i,t}^*)]^2 \quad (A2.3)$$

The difference between actual and fitted paths in the squared brackets can be interpreted as the conditional or unexplained component of property crime and it is used in the distribution dynamics analysis.

Notes

1. Data on property crime rates are available at <http://www.ucrdatatool.gov/>.
2. For a more detailed description of the UCR program, in particular, and of crime reports and statistics, in general, see Rennison (2009).
3. Data on personal per capita income are available at http://www.bea.gov/iTable/index_regional.cfm.
4. The time series of the CPI index is available at <http://www.bls.gov/cpi/>.
5. The number of state police employees is reported in the annual publications of the Uniform Crime Reports: to the best of our knowledge, for the earliest periods only the paper version is available (at <https://archive.org/index.php>). There are missing values for some states in few years: in the period from 1968 to 2013, 1.3% of the data are missing. To overcome this problem, in this study, the missing values have been imputed in two ways: by replacing them with the average of the values of the two closest years or, if the missing value is at the end of the period considered, by forecasting it using an ARIMA model. Moreover, in order to have a coherent measure of state police employees, only the police agencies that report their employment since the beginning of the period analysed have been considered.
6. This choice is motivated by a lack of observations for the New York state in the period 1960-1964 and by the use of out-of-sample lags and leads in the next conditioning procedure.
7. Therefore, the state of Alaska and the Hawaii are excluded from the sample, as well as the District of Columbia.

8. The bimodal shape is confirmed by the formal Silverman's (1981) test performed on the final distribution, whose shape is very close to that of the ergodic distribution.
9. In states with high crime rates there might be less investments and, hence, lower income. Moreover, the higher the crime rate is, the more a state should be willing to invest in police and crime repression.

Acknowledgments

I am grateful to Stefano Magrini, Paolo Pellizzari and three anonymous referees for their extremely useful comments and suggestions. I also thank Margherita Gerolimetto and Stefano Magrini for the code used to implement the kernel density estimator.

References

- Baltagi, B. H. (2006). "Estimating an Economic Model of Crime Using Panel Data from North Carolina." *Journal of Applied Econometrics*, 21(4): 543–547.
- Barro, R. J. (1991). "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics*, 106(2): 407–443.
- Barro, R. J., and X. Sala-i-Martin (1991). "Convergence Across States and Regions." *Brooking Papers on Economic Activity*, 1991(1): 107–182.
- Barro, R. J., and X. Sala-i-Martin (1992). "Convergence." *Journal of Political Economy*, 100(2): 223–251.
- Baumol, W. J. (1986). "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show." *American Economic Review*, 76(5): 1072–1085.
- Becker, G. S. (1968). "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(2): 169–217.
- Cohen, L. E., and M. Felson (1979). "Social Change and Crime Rate Trends: A Routine Activities Approach." *American Sociological Review*, 44(4): 588–608.
- Cook, J., and S. Cook (2011). "Are US Crime Rates Really Unit Root Processes?" *Journal of Quantitative Criminology*, 27(3): 299–314.
- Cook, S., and D. Watson (2013). "Breaks and Convergence in U.S. Regional Crime Rates: Analysis of Their Presence and Implications." *Social Sciences*, 2(3): 180–190.

- Cook, S., and T. Winfield (2013). "Crime across the States: Are US Crime Rates Converging?" *Urban Studies*, 50(9): 1724-1741.
- Cornwell, C., and W. N. Trumbull (1994). "Estimating the Economic Model of Crime with Panel Data." *Review of Economics and Statistics*, 76(2): 360-366.
- Donohue, J. J., and S. D. Levitt (2001). "The Impact of Legalized Abortion on Crime." *Quarterly Journal of Economics*, 116(2): 379-420.
- Gerolimetto, M., and S. Magrini (2014). "A Novel Look at Long-Run Convergence Dynamics in the United States." *International Regional Science Review*: 1-29.
- Gould, E. D., B. A. Weinberg and D. B. Mustard (2002). "Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997." *Review of Economics and Statistics*, 84(1): 45-61.
- Greenberg, D. F. (2001). "Time Series Analysis of Crime Rates." *Journal of Quantitative Criminology*, 17(4): 291-327.
- Hyndman, R. J. (1996). "Computing and Graphing Highest Density Regions." *American Statistician*, 50(2): 120-126.
- Hyndman, R. J., D. M. Bashtannyk and G. K. Grunwald (1996). "Estimating and Visualizing Conditional Densities." *Journal of Computational and Graphical Statistics*, 5(4): 315-336.
- Kapuscinski, C. A., J. Braithwaite, and B. Chapman (1998). "Unemployment and Crime: Toward Resolving the Paradox." *Journal of Quantitative Criminology*, 14(3): 215-243.
- LaFree, G. (2005). "Evidence for Elite Convergence in Cross-National Homicide Victimization Trends, 1956 to 2000." *Sociology Quarterly*, 46(1): 191-211.
- Levitt, S. D. (2002). "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply." *American Economic Review*, 92(4): 1244-1250.
- Levitt, S. D. (2004). "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not." *Journal of Economic Perspectives*, 18(1): 163-190.
- Loader, C. R. (1999). *Local Regression and Likelihood*. New York: Springer.
- Machin, S., and C. Meghir (2004). "Crime and Economic Incentives." *Journal of Human Resources*, 39(4): 958-979.
- Marvell, T. B., and C. E. Moody (1996). "Specification Problems, Police Levels, and Crime Rates." *Criminology*, 34(4): 609-646.
- Marvell, T. B., and C. E. Moody (2001). "The Lethal Effects of Three-Strikes Laws." *Journal of Legal Studies*, 30(1): 89-106.
- Paternoster, R., and S. D. Bushway (2001). "Theoretical and Empirical Work on the Relationship Between Unemployment and Crime." *Journal of Quantitative Criminology*, 17(4): 391-407.

- Quah, D. T. (1993a). "Galton's Fallacy and Tests of the Convergence Hypothesis." *Scandinavian Journal of Economics*, 95(4): 427–443.
- Quah, D. T. (1993b). "Empirical Cross-Section Dynamics in Economic Growth." *European Economic Review*, 37(2-3): 426–434.
- Quah, D. T. (1996a). "Empirics for Economic Growth and Convergence." *European Economic Review*, 40(6): 1353–1375.
- Quah, D. T. (1996b). "Convergence Empirics across Economies with (Some) Capital Mobility." *Journal of Economic Growth*, 1(1): 95–124.
- Quah, D. T. (1997). "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs." *Journal of Economic Growth*, 2(1): 27–59.
- Raphael, S., and R. Winter-Ebmer (2001). "Identifying the Effect of Unemployment on Crime." *Journal of Law and Economics*, 44(1): 259–283.
- Rennison, C. M. (2009). "Crime Reports and Statistics." In Miller, J. (Eds.), *21st Century Criminology: A Reference Handbook*, Thousand Oaks: SAGE Publications: 383–391.
- Rosenfeld, R., and R. Fornago (2007). "The Impact of Economic Conditions on Robbery and Property Crime: the Role of Consumer Sentiment." *Criminology*, 45(4): 735–769.
- Sala-i-Martin, X. (1996). "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence." *European Economic Review*, 40(6): 1325–1352.
- Shoesmith, G. L. (2010). "Four Factors that Explains both the Raise and Fall of US Crime, 1970–2003." *Applied Economics*, 42(23): 2957–2973.
- Silverman, B. W. (1981). "Using Kernel Density Estimates to Investigate Multimodality." *Journal of the Royal Statistical Society, Series B*, 43(1): 97–99.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall.
- Smith, K. B. (1997). "Explaining Variation in State-Level Homicide Rates: Does Crime Policy Pay?" *Journal of Politics*, 59(2): 350–367.

Figures

Figure 1: Time series of the main variables of the model in the presence of different income shocks.

Figure 2: Time series of the property crime rate in the US from 1971 to 2010.

Figure 3: Distribution dynamics analysis of property crime rates for the period 1971-2010 using the 48 continental US states.

Figure 4: Distribution dynamics analysis of property crime rates for the period 1971-1980 using the 48 continental US states.

Figure 5: Distribution dynamics analysis of property crime rates for the period 1981-2010 using the 48 continental US states.

Figure 6: Distribution dynamics analysis of property crime rates for the period 1971-1980 using the 9 US regions.

Figure 7: Distribution dynamics analysis of property crime rates for the period 1981-2010 using the 9 US regions.

Figure 8: Distribution dynamics analysis of real per capita personal income for the period 1981-2010 using the 48 continental US states.

Figure 9: Distribution dynamics analysis of state police employees per 100,000 inhabitants for the period 1981-2010 using the 48 continental US states.

Figure 10: Distribution dynamics analysis of property crime rates for the period 1971-2010 conditioning on income and police.

Figure 11: Distribution dynamics analysis of property crime rates for the period 1981-2010 conditioning on income and police.