Hopeless Life:

Homicide in Minas Gerais, Rio de Janeiro and São Paulo: 1981 to 1997[1]

Mônica Viegas Andrade[2]

Marcos de Barros Lisboa[3]

Resumo

Esse trabalho tem como objetivo estudar o comportamento da taxa de homicídio na população masculina e sua relação com variáveis econômicas nos estados de Minas Gerais, Rio de Janeiro e São Paulo entre 1981 e 1997. Nossa abordagem se diferencia do tratamento usual da literatura pela construção de taxas de homicídio específicas para cada idade entre 15 e 40 anos. As variáveis econômicas apresentam coeficientes significativamente diferente de zero para a população entre 15 e 19 anos. Como esperado, um aumento do salário real e uma queda da desigualdade reduzem a taxa de homicídio. Surpreendentemente, uma queda do desemprego parece aumentar a taxa de homicídio. A maior parte dos coeficientes, porém, converge para zero com o aumento da idade, tornando-se não significativos a partir dos 20 anos. Além disso, identificamos a existência de inércia nas taxas de homicídio: gerações com maior taxa de homicídio quando jovem tendem a apresentar maiores taxas de homicídio durante todo o restante do seu ciclo de vida. Dessa forma, se as variáveis econômicas induzem uma alta taxa de homicídio entre os jovens em determinado ano, essa taxa tende a permanecer elevada para a geração durante seu ciclo de vida independente do comportamento posterior da economia. Utilizamos, nessa análise, uma reformulação do tradicional modelo Logit que incorpora a variável dependente defasada.

Abstract

This paper studies the male homicide rate and its relation to economic variables in the states of Minas Gerais, São Paulo e Rio de Janeiro between 1981 and 1997. The novelty of our approach is the construction of homicide rates specific for each age between 15 and 40 years of age. The coefficients for the economic variables are significantly different from zero for the population between 15 and 19 years of age. As expected, an increase in real wage and a decrease in inequality reduce the rate of homicide. Surprisingly, a decrease in the unemployment rate seems to increase the rate of homicide. Most coefficients, however, converge to zero as a generation gets older, becoming non-significant for the population aged 20 years old or more. We also identify an inertia component in the homicide rate: generations with higher homicide rates when young also tend to have higher homicide rates over the remainder of their life cycle. Therefore, if economic variables induce a high rate of homicide among young people in a certain year, this high rate tends to persist over the generation's life cycle independent of the economy's later behavior. Regressions are performed using a reformulation of the standard Logit model that incorporates a lagged dependent variable.

Key Words: Homicide, 4, Crime and Economy.
Code J.E.L.: K42, I12.

---

[1] Naércio Aquino and Carlos Martins were kind enough to comment on a preliminary version of this article. Unfortunately, we are the only ones responsible for any remaining errors.

[2] Escola de Pós-Graduação em Economia da Fundação Getúlio Vargas-EPGE/FGV/RJ e Departamento de Economia- Face/Cedeplar/UFMG.

I. *Introduction*

In the last twenty years, violence has grown at a frightening rate in the states of Rio de Janeiro and São Paulo, becoming, indeed, the main cause of death for men between 15 and 44 years of age. In São Paulo, the number of homicides per 100,000 inhabitants went from 54.4 in 1981 to 128.4 in 1995 for young men between 15 and 24 years of age, and from 49.3 to 106.2 for men between 25 and 44 years of age.[4] This increase represents a growth of respectively 136% and 115%. Rio de Janeiro reveals an even larger homicide rate per 100.000 inhabitants for this period, though the increase in violence has not been as marked as in São Paulo. For men between 15 and 24 years of age, the mortality rate went from 148.9 to 275.3 between 1981 and 1995, representing a growth of 85%.

In Andrade and Lisboa (2000), we calculated the years of life lost to different mortality causes. The figure estimates how many additional years an individual would have lived, on average, if the cause of his death did not exist. The surprising thing about the result is the relative growth of violence as a cause of mortality, while most of the other causes show a tendency to converge towards the figures of developed countries, though these remain higher. In the beginning of the nineties, violence emerged as the main cause of the reduction in years of life for males in the state of Rio de Janeiro, and as culprit number two in São Paulo, being surpassed only by child mortality. Therefore, between for instance 1981 and 1995, the average number of lost years of life for each male residing in the state of Rio de Janeiro went from 1.57 to 3.42. This means that if the problem of homicide were to be eradicated, each man would live an additional 3.42 years. The high homicide rate in the male group lead us to conduct an investigation in greater detail about the history of its development and its possible relationship to economical factors.

The objective of this work is to analyze the evolution of the homicide rate in the states of Minas Gerais, Rio de Janeiro and São Paulo between 1981 and 1997. We tried to verify, in particular, the existence of a relationship between this evolution and economic

---

[3] Escola de Pós-Graduação em Economia da Fundação Getúlio Vargas- EPGE/FGV/RJ.
[4] This information can be found in Andrade and Lisboa (2000).

variables such as real wage, unemployment, and Gini**,** among others. How intimately connected is the growth of the homicide rate to the economic crisis of the eighties and nineties? Is it possible to identify a recurring pattern in the homicide data?

The relationship between economic variables and violence has been the object of several works of applied economy, which often use the rate of homicide per 100.000 inhabitants as a measure of violence. However, this method of measurement does not seem adequate to us, for three main reasons. Firstly, the homicide rate varies considerably between ages and sexes. Generally, the main victims of homicide are young men between 15 and 30 years of age. Thus, the homicide rates per 100.000 inhabitants can vary significantly between regions due only to differences in age and sex distribution. For example, the frequency of death among males between 15 and 24 was about three times as high in 1981 in Rio de Janeiro as in São Paulo (149.8 and 54.4, respectively), while the homicide rate for 100.000 inhabitants in the same period was only twice as high (30.64 and 15.31). This result reflects the superior average age and greater participation of women in the population of the state of Rio de Janeiro, in contrast to the state of São Paulo. This typical case of omission of variables causes a difference in magnitude that may results in distortions of controlled homicide rate development analyses.

Secondly, the relationship between homicide rates and economic variables can shift along a life cycle. Several legal activities show increasing returns with the specialization that results from long time practice. Even if there are no increasing returns, the previous participation of the worker in the labor market can be a sign of his quality or productivity to employers. Besides, past involvement in illegal activities may compromise access to legal activity. In this way, individuals with experience in the legal labor market can have better access to remuneration and employment than those who practice illegal activities, and this difference may grow along the life cycle.

Thus, it is possible that the impact of economic variables on the portion of the population that are involved in illegal activities may vary according to the life cycle stage. For a youth, the difference between legal and illegal activities may be smaller than for an older male who participates in the legal labor market. In this case, the economic variables may have a greater impact on the young than on the old. If the rate of homicides grows along with the portion of the population that is involved in illegal activities, the impact of the economic variables on the homicide rate will also shift along the life cycle.

In third place, the differentiated access to the labor market among the population that participates in the legal labor market and those who participate in illegal activities may result in generation effects on the homicide rate (inertia effect). Let us suppose once more that the homicides grow with the portion of the population involved in illegal activities. If the return to legal activities is very hard, then a generation which, when young, is in its majority dedicated to illegal activities, tends to maintain these proportions during the whole life cycle and, thus, also a higher rate of homicide. This means that an inertia effect of the homicides may exist in each generation. And especially, the homicide rate in one year may be high not because the economic variables that year are behaving in a certain manner, but because this behavior was observed when the generation was young, and its effect are diluted along the life cycle of that generation.

The existence of the inertia effect can be tested by separating the homicide data for each age group and following this data for each generation for several years. The existence of correlation in the homicide rates per generation, when controlled by the remaining variables, is an indicator of the possible existence of this inertia effect. In this case, the homicide rate of a given generation in a given year would be one of the relevant variables to predict the rate of the same generation for the following year. A generation that was violent when young would tend to present greater rates of violence along their whole life cycle.

In this work, violence is measured through the arrangement of death rates by homicide for each age, sex, year and residential area. The death rates are arranged from homicide data divided by resident population of each region and according to age, sex and year. The arrangement of this data base allows us to estimate the relationship between probability of death by homicide and economic cycles for each specific age. As the phenomenon of violence is concentrated in the male population of active age, we calculated the probability of death for men from 15 to 40 years of age. This treatment of the data also allowed us to construct the data base according to *cohorts*.[5] Each cohort was defined by the year in which the men were 15 years old.[6]

There are basically three forms of empirical approach that try to explain the causality between violence and socio-economic conditions: cross-section analyses, time-series analyses and work based on victimization research that follows individual behavior. The results observed in these three types of analysis are quite varied, particularly where unemployment impact is concerned. These results are discussed in the fifth section and compared with those obtained by our analysis.[7]

In this work we used a mix of cross-sectional and time-series analysis. We used death rates by homicide for three Brazilian states; Rio de Janeiro, São Paulo and Minas Gerais, during the period from 1981 to 1997, with annual data. The time-series analysis has the advantage of being suitable for the study of relationships between economical cycles and violence, and for allowing us to study the inertia effect. In fact, one of the problems with the cross-section analysis is the possible occurrence of spurious correlation. The possibility of migration between states suggests that similar individuals in distinct regions have a similar quality of life; otherwise they could simply emigrate to regions with a higher quality of life. For this reason, regions with higher homicide rates must offer inhabitants compensations, such as higher wages or better access to other public commodities besides public safety. In that case, the existence of a positive correlation in a cross-section data base, for instance, between real wage and violence, may not mean that increases in real wage result in more violence, but just that in a given region, the decline in life-quality due to greater violence is compensated by a larger real wage, which reduces migration. It is also possible that a simultaneous increase in real wage throughout the several regions might result in a reduction of violence. The construction panel data base may allow the

[5] The concept of *cohort* is based on the definition proposed by Ryder (1965). Each *cohort* is defined as a group of individuals who experience the same events during the same period of time.
[6] Few works in the existing literature analyse the cohort data in relation to the lack of availability of this type of research. Tauchen and Witte (1994), Tauchen, Witte and Griesger (1994) and Steffensmeir (1992) analyse data in cohorts.
[7] Freeman (1994) summarises the results found for American economy with the three different approaches.

observation of parallel fluctuations of the variables that might explain the homicide rate in all regions, which would avoid spurious correlation. We will get back to this issue in the fifth section, when comparing our results to those of existing literature.

The estimation method used in this work is a generalization of Berckson's Minimum Chi-Square Method.[8] This method consists in the estimate of a logit model for qualitative variables when the data is available in the form of frequencies. In the specific case of this work, the data is grouped according to age, region of residence and year of occurrence of homicide. The dependent variable, obtained through the Datasus mortality information system, is the probability of death by homicide for each age, region and year. Socio-economic indicators for each region as well as specific generational attributes were used as independent variables.

In order to examine the possibility of the inertia effect, we used the lagged death rate by homicide rate for each cohort as a independent variable in each year. However, this treatment implies in the alteration of the variance formula of the traditional Berckson model, and requires a different correction of heterocedasticity**.** The method and the estimated model are presented in detail in the methodological section.

This work presents three main results. Firstly, the organization of data according to cohorts of individuals appears to be an adequate approach for the understanding of the violence cycles. The observation of death-probability development according to each cohort described in graphs1, 2, 3, and 4, supports the hypothesis of persistent cycles of violence. The pattern of behavior between cohorts is practically the same: at first, death probability increases with age, reaching a peak, in most cases, between 20 and 25 years of age, after which it starts declining.[9] Each cohort's cycle is of approximately 25 years. This result is pronounced in Rio de Janeiro, but the pattern repeats itself in the other two states as well. For the state of Rio de Janeiro, two graphs are shown. The first displays the

[8] See Maddala (1983) and Amemya (1985).

[9] The relationship between rate of participation in crime and age is noted by several authors and appears sound for any kind of crime. Grogger (1997) justifies this pattern through the behavior of salaries that grow with age as individuals acquire more experience. If the behavior of crime rates is sensitive to that of the

evolution of the cohorts that were 15 years of age in 1982, 1983 and 1984 respectively. Since the available data base only allowed us to obtain information from years previous to 1997, the maximum amount of time we were able to follow these cohorts is 16 years. The second graph shows the cohorts that were 15 years old in the period between 1972 and 1976, demonstrating the rhythm of decrease in death probability with age. The arrangement of probability data per age seems to suggest that the homicide rate each year is the sum of superimposed waves of violence, each corresponding to a specific generation.

The second result of the research refers to the relevance of the economic variables in explaining the homicide rate. The economic variables may be relevant for youths between 15 and 19. After 20 years of age, the most important variable to explain violence is the inertia component, measured in this work through the inclusion of lagged probability. This result differs from existing works in the field by incorporating the cohort effect. The lagged probability is the probability of an individual from a given cohort to have been murdered the previous year.

A possible interpretation of this result relates, as we have already suggested, the generation age and the population portion dedicated to illegal activities. Younger individuals would move between legal and illegal activities with a greater facility than the older, and this movement would be influenced by economic variables. The increase of the population dedicated to illegal activities, in turn, would result in greater rates of homicide. Thus, a generation whose members, when young, are in great numbers involved with illegal activities, tends to maintain this high number during its whole life cycle and, thus, maintains a greater homicide rate as well. This hypothesis, however, is still to be tested. What remains as a result of this analysis is that public safety policy should perhaps focus on the young male population, and that the homicide rate in this group impacts on the homicide rate of the whole generation. However, the effects of such a policy can only be perceived in the long run.

---

salaries, then when wages go up the crime rate decreases, since the cost of opportunities in criminal activities grows.

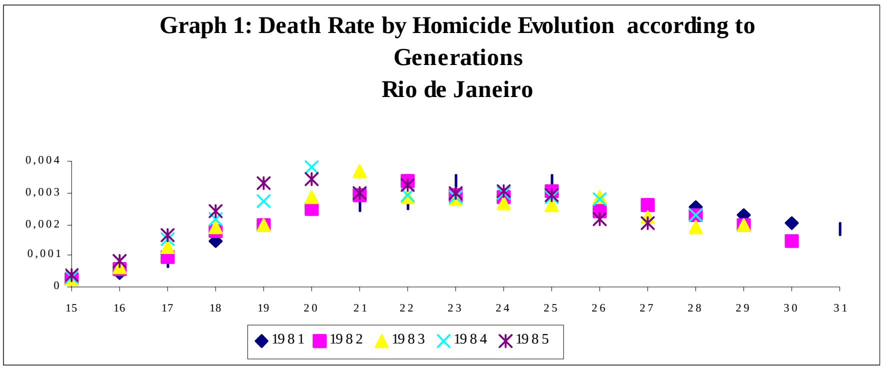
Thirdly, the difference between the homicide rates in Rio de Janeiro and São Paulo are significantly reduced when these rates are controlled by the inertia effect and by the economic variables. For most age groups, the non-controlled difference of the homicide rates is about twice as large in Rio de Janeiro as in São Paulo, while in most regressions this difference is reduced by about 20 %.

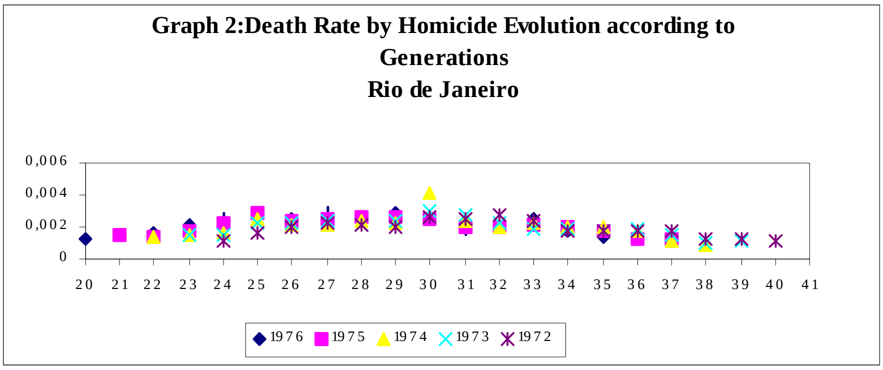
Probably, the main frailty of our analysis is the non-inclusion of a variable as proxy for the penalty system and the safety policies. The only available data for a continuous series are the police-force occurrences, which are correlated with the criminality rate itself.[10] Unfortunately, we were unable to obtain a reliable data base describing the number of arrests per year, public safety spending, or any other indicator of police activity. This lack of minimally reliable data describing any aspect of the public safety policy is perhaps the largest problem with empirical analysis of violence in Brazil. It also reflects, in our view, an apparent carelessness of the state concerning any kind of long-term public safety policy. We were particularly surprised to find that there is no time-series analysis for the period examined that states the number of prisoners in each one of the three states.

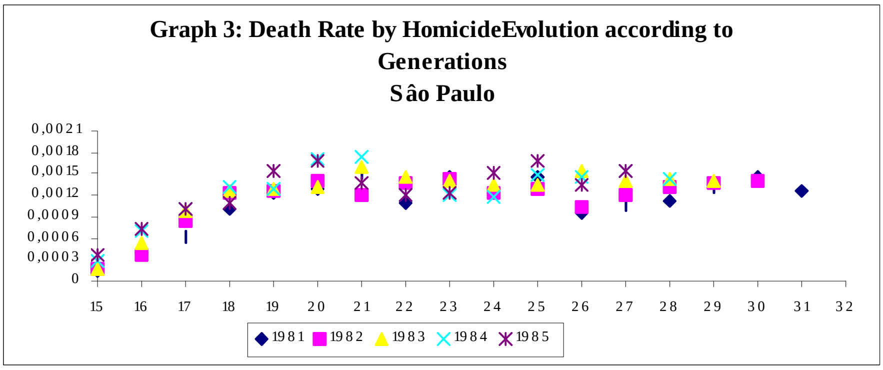
Our results seem to suggest that eventual differences in public safety policies may not have such a significant impact on the difference between the homicide rates observed in Rio de Janeiro and São Paulo as suggested by the non-controlled data. These results, however, must be treated with care. Public safety policies may be correlated to the economic variables of each state. More affluent states would have greater resources to put a better public safety policy into operation. Also, the economic variables used are not specific to each cohort, but pertain to all cohorts located in the same region each year. Therefore, these results must be seen as a preliminary outcome, based on which additional study about the dramatic increase in the deaths by homicide during the eighties and nineties should be conducted.

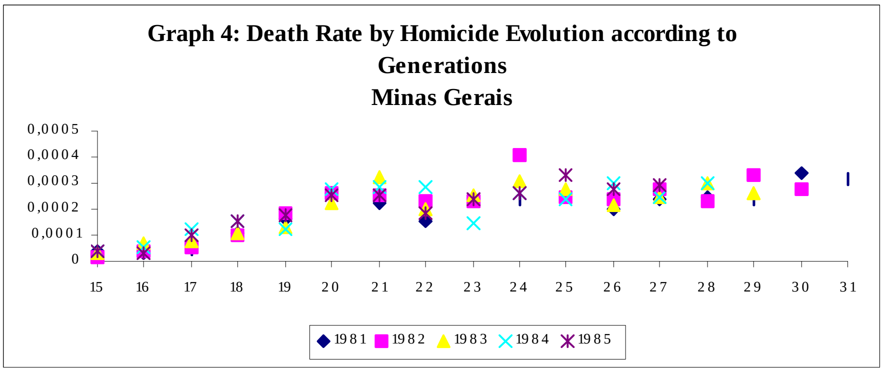
[10] In a previous version we attempted to include dummy variables associated to elections as approximations to the public safety system. However, the number of age group observations in the data bases is insufficient for a significant estimate of the coefficients.

This work contains five additional sections. The next section describes the data base used and the main variables included in the model. The third section discusses the methodology. The fourth section shows the main results found in the model estimate. The fifth section discusses some figures in the existing literature as well as estimates made for the Brazilian case, confronting them with the empirical results found in this work. The sixth section discusses the possible extensions of this work.


Graph 1: Death Rate by Homicide Evolution according to Generations
Rio de Janeiro
0,004
0,003
0,002
0,001
0
15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
1981 1982 1983 1984 1985


Graph 2:Death Rate by Homicide Evolution according to Generations
Rio de Janeiro
0,006
0,004
0,002
0
20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41
1976 1975 1974 1973 1972


Graph 3: Death Rate by HomicideEvolution according to Generations
Sâo Paulo
0,0021
0,0018
0,0015
0,0012
0,0009
0,0006
0,0003
0
15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
1981 1982 1983 1984 1985


Graph 4: Death Rate by Homicide Evolution according to Generations
Minas Gerais
0,0005
0,0004
0,0003
0,0002
0,0001
0
15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
1981 1982 1983 1984 1985

II The data base

II.2 Series used

In this work, the variable utilized to measure the crime rate is the death rate by homicide. For each year and region we divided the total number of homicides for each age by the male resident population of the same age. The population data was obtained from the 1980 and 1991 censuses and the 1996 population count. For the other years, we estimated the population by using a log-linear interpolation.

The data base used for the death rate data was the Death Rate Information System (Sistema de Informação de Mortalidade, or SIM) for the period of 1979 to 1997[11], made available by the Ministry of Health through the National Foundation of Health and Datasus. The primary data source of this base are the death certificates emitted by public notaries. This data base is extremely rich and contains information about dates of decease, age, sex, marital status, place of occurrence, cause of death, area, neighborhood and municipal district of residence, occupation and education. In spite of the great range of information, this data is poorly filled out concerning variables such as education, marital status, occupation, among others, which limits its usefulness. In this work we used only the variables prioritized by the Ministry of Health; age, sex and cause of death, for which the blanks are on average 7%. Until 1996, death cause was codified according to the 9th International Illness Classification, CID9, and in 1997 according to the 10th revision, CID10.

In the states of São Paulo and Minas Gerais, the number of homicides was computed by using the observed rates. However, in the state of Rio de Janeiro, two death causes were used to calculate the number of homicides: homicides and injuries caused intentionally and then some causes classified as ‘other violent actions’ . Rio de Janeiro shows an elevated number of deaths caused by fire arms and other weapons but classified as ‘of unknown intent’. While such deaths in São Paulo, for instance, never go further than

[11] Though the death rates are available from 1979 on, the two initial years of the research present many data processing problems.

7% of the death per homicide rate of men between 15 and 24 years of age, in Rio de Janeiro there are years when this number is 37%, though it seems to behave rather erratically.[12] These results suggest a problem with the death certificate data basis in that state.[13] To get around this obstacle, we made the following adjustment to the homicide data: for each year and for each age group, we calculated the percentage of deaths caused by weapons and of unknown intent against the total of homicides in São Paulo. This percentage was then used to re-compute the total homicides in Rio de Janeiro for the same year and age group. The basic presumption for this procedure is that the deaths caused by fire arms or other weapons, though unintentional or at least hard to classify as intentional, must be similar in the various regions.

The choice of the states of São Paulo, Minas Gerais and Rio de Janeiro for analysis is due to two motives: firstly, as has been mentioned in the introduction, the labor market integration of these three states, and secondly, the lesser rates of sub-enumeration of deaths. In the north and northeast regions, for example, the sub-enumeration rates estimated by IBGE reach 50% in some states.

The computing of the crime rate through death probability at each age allowed us to construct a data base in two ways. In the first case we arranged a base for each age group and each state, from 15 to 40 years of age, which corresponds to the age group in which violence is the main cause of death. The data base for each specific age contains in average 42 cells. In this way, we obtained 25 data bases and could estimate the specific coefficients for each age.

The second data base was constructed by observing individuals according to their cohort, which is defined by the year in which the individuals were 15 years of age. The first cohort corresponds to the individuals who were 15 in 1981. The first observation of this cohort is the probability of death by homicide of 15-year-old males in 81, the second observation is the probability of death of 16-year-old males in 82, the third, probability of 17-year-old males in 83 and so on, until the probability of death of 31-year-old males in

[12] For the time series of this data, see Andrade and Lisboa (2000)
[13] This problem is well known in the existing literature. See for instance Carneiro and Phebo (1999)

1997. The second cohort corresponds to 15-year-old males in 1982, and so on. In this way we constructed 43 cohorts for each state, computed to 1410 cells. The organization of the observations according to cohorts allowed us to perceive the inertia effect in the violence cycle. Furthermore, the lagged probability rate was constructed by incorporating the cohort effect, since it denotes the probability of an individual of that cohort to have died in the previous year. Graphs 1 and 4 clearly illustrate the overlapping of the violence cycles. Each new cohort is slightly displaced, so that each generation shows different results from the previous.

The tradeoff for constructing these two data bases is worth noting. On one hand, the bases for each age group allows inference about the differentiated impact of the several independent variables on the probability of homicide. Specifically, the degree to which these coefficients change significantly with age can be verified. This approach, however, reduces the number of cells. The complete data base, though it allows for results of greater significance, requires the specification of a previous functional form relating the various coefficients of independent variables to the age of the generation.[14]

When possible, we attempted to incorporate the variables discussed in the existing literature about the economy of crime into the control variables used in this work, following Becker's contribution (1968). Due to blanks in the death rate data base used, we could not include the variables of the individuals themselves, so as not to distort the sample selection. To get around this problem, we used average variables for each region and year.

The control variables included were: average educational level of the economically active population, Gini coefficient, unemployment rates, percent of households commanded by women, real average wage of the employed population, price levels, lagged probability and two dummy variables, one for the state of Rio de Janeiro and one for the state of São Paulo.[15] These are the control variables commonly used in empirical works. The level of education and the real wage are measures of the gains from

[14] A possible extension of this work would be to utilize a non-parametric approach to estimate this functional form.

legal activity; the unemployment rate is tied to the opportunities of individuals in the labor market; the rate of households commanded by women is an approximation of the degree of social changes and integration. The index of income inequality, measured by the Gini coefficient, describes the relative position of individuals in society. For poor individuals, a rise in inequality makes for a greater difference between gains from criminal activity and gains from legal activity. All these variables are arranged according to the year and region.

The inflation rate was included in order to catch possible distortions of relative prices due to the inflationary process in these two decades, which may alter income from activities, aside from the possible implications of the greater instability of real income in periods of high inflation. Unlike other variables, the rate of inflation is only listed by year, being the same in all three regions.

The dummies for the states of Rio and São Paulo were included in order to verify the existence of regional specifications not measured by the remaining control variables. The lagged probability functions as a gauge of the inertia effect.[16] It is arranged according to generation and region. Consider, for instance, the generation that was 15 years of age in 1985 in Rio de Janeiro. In 1990, when these individuals are 20 years old, the lagged probability is the death probability by homicide of the individual who were 19 years old in 1989. For each year, we used the homicides rate for this same generation in Rio de Janeiro the previous year as one of the control variables.

It should be noted that, unfortunately, as we have already discussed in the introduction, we did not use any indicator of the public safety policies, which may have impacted crime rates and which may be different in the several states along the whole period. We were unable to find any sort of reliable data base that describes the number of arrests for each region, or even the total number of existing prisoners. The only available data was the total number of occurrences, which, as it is correlated to the criminal activities themselves, does not serve as indicator of public safety policies.

---

[15] The lagged probability data was also constructed for cohorts. Witt and Witt (1998) show in a time series model that the increased female participation in the labor market is positively correlated to higher crime rates.

[16] Fanzylber, Loayza and Lenderman (1999) also use the lagged crime rate as a control mechanism.

II.2 Behavior of the variables.

The non-controlled homicide rates show a distinct behavior in the states of Rio de Janeiro and São Paulo in contrast to the state of Minas Gerais. In the first two, the rates show a relatively similar behavior for the various age groups. In the state of Rio de Janeiro the behavior is rather clear. With the exception of the 35-age group, all groups show a growth in death probability, peaking in 1989, followed by a long slant from 90 to 93, a small elevation in 94 and a tendency to decline from 1995 on. In the case of the 35-age group, growth starts in 1989.

In São Paulo the movement is less homogeneous, presenting a distinct pattern for the younger groups. The death probability for the 15, 18 and 21 age groups shows similar trajectories: growth until 87, a dip in 1988, new ascending cycle in 90 and 91, decline in 92 and 93 and a tendency towards increase from 1994 on. As for the 25 and 30 age groups, there is growth until 1983 followed by stability until 1986, a small peak in 1987, decline again in 88, new cycle of growth in 89 and 91, retraction from 91 to 93 and a tendency to increase from 1994 on. The contrast between the two states suggests a more cyclical behavior in São Paulo and a tendency towards increased violence at the end of the 90's, contrary to Rio de Janeiro where, for all ages, the death probability seems to indicate a tendency to decline.

The state of Minas Gerais presents quite inferior homicide rates in relation to the other two states. Aside from this, there does not seem to be a common pattern of behavior along time for the non-controlled rates of the various ages.

The development of the variables associated with individuals, level of education and percent of households commanded by women shows a linear growth along the two decades, which is practically homogeneous between the three studied regions. In Brazil, the average level of education of the economically active population increases one year of study per decade. Due to this behavior, which is virtually constant along time, these variables show a significant correlation with any tendency variable.[17] The same correlation pattern can be perceived for the Gini coefficient, which presents little change along time, showing differences only in level between the regions. Such a pattern makes this variable highly correlated to the region dummies.

The remaining economic variables show cyclical behavior along the two decades. The unemployment rate increases significantly with the recession starting in 1981, but fall in 82 and 83, when a fallback of the GDP can still be observed. The state of Rio de Janeiro presents significantly smaller alterations than São Paulo in unemployment rate practically during the whole period after 1986. This behavior is probably associated to the growth of the informal sector of the economy. The behavior of the average real wage of the employed population is quite similar between the three states, apparently reflecting GDP cycles: the real wage is drastically reduced from 1981 to 1984, recovers from 1985 on and increases with the coming of the C*ruzado plan*, stabilizing and then decreasing again from 90 to 1993. With the inflationary control, it begins to show a tendency of growth from 1994[18].

[17] The correlation matrix for all variables is shown in the table appendix

[18] A whole version of this article that contains a methodological appendix and all detailed results is available at www.fgv.br/epge/home/publi/

Graph 5: Death Rate by Homicide Evolution according to Age
Rio de Janeiro

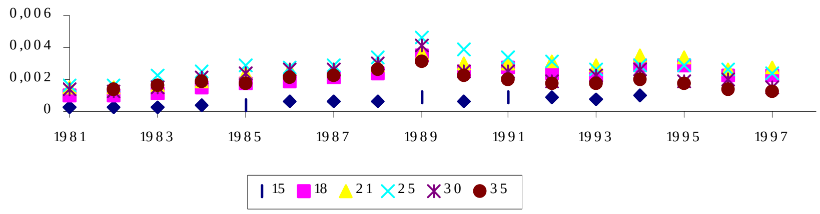
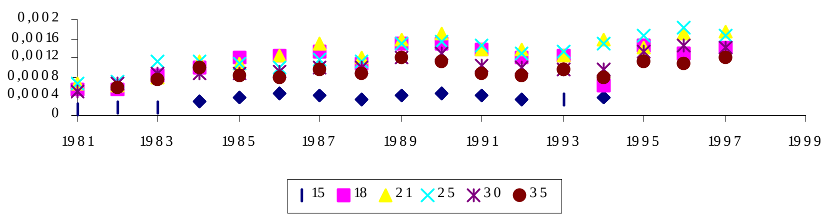
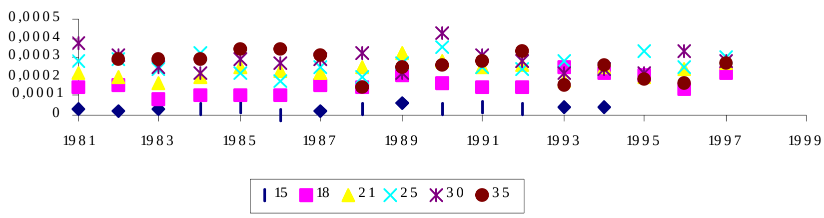


Gráfico 6: Evolução da Taxa de Homicídios por Idade
São Paulo



Graph 7: Death Rate by Homicide Evolution according to Age
Minas Gerais



III. Methodology

III.1. Basic Model

The estimation method used in this work is a generalization of the Minimum Chi-Square method applied to the Logit model, also known as the Berckson Model. Logit models are used when the dependent variables are qualitative, represented by binary variables: 1 if the event occurs, and 0 if the event does not occur. In this model, it is assumed that the probability of occurrence depends on the independent variables according to the following functional form: $P_i = P(y_i = 1) = \Lambda(X'b) = \frac{e^{x'b}}{1+e^{x'b}}$.

In the case of a multiple observation model we can apply the Logit model with grouped variables. In this model we have, for each age, n_i observations, which corresponds to the total number of males of a specific age resident in each region, and the event "death by homicide" occurs for m_i observations according to a binomial distribution.[19]

Let p_i be the probability of m_i occurrences of death by homicide in n_i observations, and $\hat{p}_i$ the frequency observed**.** As usual, we may write,

$$\begin{aligned} Log\frac{p_i}{(1-p_i)} &= b'x_i \\ Log\frac{\hat{p}_i}{(1-\hat{p}_i)} &= b'x_i + u_i \end{aligned} \qquad (1)$$

where u is an arbitrary variable with a zero-average . Thus,

$$u_i = Log\frac{\hat{p}_i}{(1-\hat{p}_i)} - Log\frac{p_i}{(1-p_i)}. \qquad (2)$$

The shortcoming of this model is that according to the hypothesis that the events follow a binomial distribution the variance is not constant, being the model heterocedastic. Using a Taylor-series expansion of $Log\frac{\hat{p}_i}{1-\hat{p}_i}$ around p_i, we get :

[19] This grouping goes beyond the simple frequency calculation observed in homicides because the probability was estimated according to adequate demographic techniques.

$$\log\frac{\hat{p}_i}{1-\hat{p}_i} \cong \log\frac{p_i}{1-p_i} + (\hat{p}_i - p_i)\left[\frac{1}{p_i(1-p_i)}\right] \quad (3).$$

Thus, we have: $\mathrm{var}(u_i) = \left[\frac{p_i(1-p_i)}{n}\frac{1}{p^2{}_i(1-p_i)^2}\right] = \frac{1}{n_i\, p_i(1-p_i)}$ estimated by

$$\mathrm{var}(u_i) = \frac{1}{n_i\,(\hat{p}_i(1-\hat{p}_i))}.$$

We can estimate the model using a weighted least square method, defining as the weight: $w = \left(\frac{1}{\mathrm{var}}\right)^{\frac{1}{2}}$.

As we included lagged probability in the estimated basic model, we need to add a new term in the variance calculation. The equation (3) must be rewritten in the following manner:

$$F^{-1}(p_{it}) = F^{-1}(\hat{p}_{it} + e_t) \cong F^{-1}(\hat{p}_{it}) + \left[\frac{dF^{-1}(\hat{p}_{it})}{d\hat{p}_{ti}}\right]e_i = F^{-1}(\hat{p}_{it}) + \frac{1}{f(\hat{p}_{it})}e_i \quad (4)$$

By including lagged probability in the model, we can now describe it like this:

$$F^{-1}(p_{it}) \cong bX'_t + g\left[bX'_{t-1} + \frac{1}{f_{t-1}(\hat{p}_{t-1})}e_{t-1}\right] + \frac{e_t}{f_t(\hat{p}_t)} \quad (5)$$

where γ is the lagged probability coefficient. Thus, we get the following variance equation:

$$\begin{aligned}
&\mathrm{var}\left[g\,\frac{e_{t-1}}{f_{t-1}(\hat{p}_{t-1})} + \frac{e_t}{f_t(\hat{p}_t)}\right] = \\
&= \frac{g^2}{f_{t-1}{}^2(\hat{p}_{t-1})}\mathrm{var}[e_{t-1}] + \frac{1}{f_t{}^2(\hat{p}_t)}\mathrm{var}[e_t] \\
&= \frac{g^2}{f_{t-1}{}^2(\hat{p}_{t-1})}\frac{\hat{p}_{t-1}(1-\hat{p}_{t-1})}{n_{t-1}} + \frac{1}{f_t{}^2(\hat{p}_t)}\frac{\hat{p}_t(1-\hat{p}_t)}{n_t}
\end{aligned} \quad (6)$$

As $f^2(\hat{p}_t) = \hat{p}^2{}_t(1-\hat{p}^2{}_t)$ the final variance formula is given by:

$$\text{var} = \frac{g^{2}}{\hat{p}_{t-1}\left(1-\hat{p}_{t-1}\right)n_{t-1}} + \frac{1}{\hat{p}_{t}\left(1-\hat{p}_{t}\right)n_{t}} \quad (7)$$

The model is estimated in two stages: first we estimate the variance, and after we estimate the explanatory variable coefficients. A simple generalization of the argument used by Amemya (1985, p. 276-277) shows that this estimator distribution converges to a asymptotic normal distribution.

III.2 Testing the hypotheses

We used two tests in this work: the model specification test and the coefficient robustness test. The two tests are based on the same statistics, the weighted sum of squared residuals. In the specification test, we classified the models according to the fitting to the observed variables. This test follows a procedure proposed by Li (1977). By denominating

$$F^{-1}\left(\hat{p}\right) = \hat{L} = \log\frac{\hat{p}_i}{1-\hat{p}_i},$$

where $\hat{p}_i$ is the observed frequency, we calculate the sum of squared residuals weighted by the estimated variance, that is:

$$\sum_{t=1}^{T}\left(\hat{L}_i - x'b_i\right)s^{-2}\left(\hat{L}_i - x'b_i\right) = c^{2}.$$

Then, for each model, we divide the chi-square statistic figure obtained by the degrees of freedom, c^{2}/gl. The degrees of freedom corresponds to T-K, where T is the number of cells and K the number of estimated parameters. In order to establish a classification of the models we observe the lower figures in the obtained statistics, as they indicate a better adjustment of the specifications of the observed values.

In order to test the coefficient significance, we implemented the test described in Amemya (1985). In this case, we calculated the sum of squared residuals weighted by the estimated variance of the complete model and by the variance estimated in the restricted models, comparing the difference with the chi-square statistics. That is, if we define the sum of the squared residuals as $SQR = \sum_{t=1}^{T} \left(\hat{L}_i - x_i' b_i\right) s_i^{-2} \left(\hat{L}_i - x_i' b_i\right)$,

we cannot reject the significance of the restricted variables as long as:

$$SQR_{\text{complete model}} - SQR_{\text{restricted model}} > \chi^2_q$$

where q is the number of restricted variables.

III.3 Estimated models

For each base arranged by age, 11 different specifications were run, including the variables: unemployment, past unemployment, real wage, Gini coefficient, two region's dummies, price index and lagged probability. Due to the correlation existing between levels of education and rate of households commanded by women with region's dummies and lagged probability, these variables were not included in the model. As has been previously mentioned, these variables show a monotone growth along time and therefore have a strong correlation with the tendency and level variables included in the model. The Gini coefficient, though also presenting high correlation particularly with the region's dummies, was included in some specifications. The specification using the past unemployment rate was included because in some economies the relationship between violence and unemployment tax is different from the relationship verified for the present unemployment rate. The following table systematizes the eleven estimated models.

Table 1: Models Specification

Model	Equation
Model 1	**Real Wage + Unemployment rate + GINI**
Model 2	**Real Wage + Unemployment rate + GINI+ Prob (-1)**
Model 3	**Real Wage + Unemployment rate + GINI+ DummyRio + DummySp**
Model 4	**Real Wage + Unemployment rate + GINI+ DummyRio + DummySp + Prob(-1)**
Model 5	**Real Wage + Unemployment rate + GINI+ DummyRio + DummySp + Prob(-1) + Ano**
Model 6	**Real Wage + Unemployment rate + DummyRio+ DummySp + Prob(-1) + Ano**
Model 7	**Real Wage + Unemployment rate + DummyRio+ DummySp + Prob(-1)**
Model 8	**Real Wage + Unemployment rate + DummyRio+ DummySp + Prob(-1) + INPC**
Model 9	**Real Wage + Unemployment(-1) + DummyRio+ DummySp + Prob(-1)**
Model 10	**Real Wage + Unemployment rate + GINI+ DummyRio + DummySp + Prob(-1) + INPC**
Model 11	**Real Wage * [(100-Des/100)] + GINI+ DummyRio + DummySp + Prob(-1) + INPC**

IV. Results

Of the eleven specifications proposed, only model 1 and 3 showed poor adjustment to the observed frequencies[20] These models distinguish themselves from the others by not including lagged probability, which certainly looks like the variable with the greatest power of explanation. For the remaining models, the chi-square statistic presents low response to alterations in the variables, thus not allowing identification of a dominant specification. This behavior is probably caused by the existence of multicolinearity between the dummy variables, Gini, year and lagged probability, as confirmed by the alterations in the coefficient when these variables are removed from the equation.

The high degree of adjustment and the similarity in the pattern of the economic variables coefficient surprised us and pointed to the robustness of the models, which may be noted by the chi-square statistic measure. This adjustment is relatively better in the age group distribution tail.

In order to analyze the coefficients, we opted for the use of models 07 and 10 as basic models.[21] The remaining models, except specifications 9 and 11, are intermediaries of these two models, distinct from model 10 through the exclusion of a few linear variables. Model 07 differs from model 10 through the exclusion of the Gini and inpc variables. As the Gini coefficient presents high correlation with regional dummies, we chose to consider model 07 as the control specification.

For the economic variables, the main result found is the evidence that consideration of the age-group variable is extremely important in the understanding of the relationship between violence and economic cycles. The behavior of economic variables, real wages and unemployment, is quite distinct both for the sign and the intensity of the relationship between the several age groups (see tables 1and 2 in appendix 1). The results show that for younger males, from 15 to 19 years old, the economic variables are important

[20]In www.fgv.br/epge/home/publi/ we have a whole version containing all results discussed here. In this version we present only the results of models 7 and 10.

[21] The significance test for the variables was only made for these two specifications and are shown in table 1 in the appendix.

to explain violence. For these age groups, the real wage shows the expected negative sign**,** suggesting that a rise in wage reduces homicides in the younger population.

Between 20 and 29 years of age the importance of the wage is virtually null. After 30, the coefficient turns positive for some age groups. The robustness tests suggest that the wage variable is significant for all ages in the control model and significant for the age groups 15 to 17, 21, 23, and some groups over 30 in the model where we included the Gini coefficient. In other words, the significance of the real wage is altered when we introduce the Gini coefficient, suggesting a correlation between these variables.

Thus, the increase in real wage has the effect of reducing the homicide rate among youths. For older individuals, however, an increase in wage may in fact increase their chance of being homicide victims. As we have already discussed, one possible explanation for this phenomenon is that youths tend to be homicide victims when they participate in illegal activities, and this participation increases as real wage decreases. Older individuals, on the other hand, tend to be victims of violence as the gains from criminal activities grows: considering the portion of the population involved in illegal activities, the increase in real wage increases the income of certain illegal activities and many of the victims of these activities are men over 30.

The behavior of the unemployment rate is quite similar to that of the real wage, with a significant and negative coefficient for youths between 15 and 20. From 21 years of age on, the unemployment coefficient is virtually zero and the robustness tests suggest that this variable ceases to be significant both in the control model and in model 10. This result is surprising due to the negative relationship obtained between unemployment and death probability, contrary to expectations. According to the Becker crime model (1968), the relationship between unemployment and economic cycles displays two effects: on one hand, the reduction of unemployment increases the opportunities of individuals in the labor market, and on the other hand, it increases the expected gains from crime. However, the latter is a secondary effect which we did not expect to dominate the relationship.

The same result appears in several time series analyses for the American economy (Freeman, 1994) and may be result of a spurious regression. The conventional labor market model establishes a functional relationship between unemployment rate, real wage and product fluctuations: the decline of the product would lead to a simultaneous decline of wage and employment, provided the impact of the product on the wealth of the individuals is not taken into account. One of the ways commonly used to investigate this possible multicolinearity is to use the lagged unemployment rate instead of the present one. The results found in this last case, however, are similar to those obtained previously. Three phenomena may be taking place: i) as mentioned, the secondary effect of the unemployment variation on the homicide rate may indeed be dominating the main effect; ii) the used series may contain a reduced number of observations, thus slanting the results, seeing as we are working only with 17 years when usually, for the American economy, time series analyses extend for an average of 40 years; iii) the unemployment elasticity seems to have altered along the decade, due mainly to the growth of the informal sector. These results are illustrated in the graphs below.

The results found for lagged probability confirm the importance of the inertia effect in explaining criminality, particularly from 20 years of age on. The lagged probability is significant and positive for all ages and initially displays a growing impact with age. From 30 on, this effect is again reduced. The robustness test as well as the specification test reveal the explanatory power of lagged probability, which seeks to incorporate the inertia effect: generations that show a high homicide rate when young tend to maintain a higher rate throughout their life cycle. Since the economic variables affect the homicide probability when the generations are young and this effect is perpetuated throughout the generation by the inertia effect, the impact of these variables is much superior to that suggested by the magnitude of the coefficients.

As discussed in the introduction, a possible interpretation of the results links the fraction of a generation engaging in illegal activities to the homicide rate of that generation in each year. Let's suppose that the return to the legal labor market of those individuals dedicated to illegal activities is costly. Let us also suppose that the homicide rate grows together with the fraction of the generation dedicated to illegal activities and that the mobility between legal and illegal activities decreases with age: that is, that youths abandon a legal activity to engage in an illegal one with greater facility than older individuals. In this case, a generation which, when young, presents a high fraction dedicated to illegal activity due to, for instance, a decline in real wage, tends to maintain this high fraction throughout its life cycle and, consequently, a high homicide rate along this same cycle.

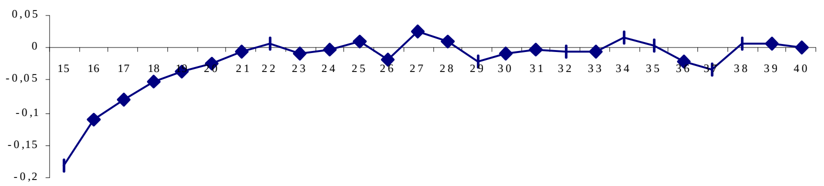
Regardless of the validity of this interpretation, however, the evidence that remains in this work is that a generation which, when young, shows high rates of homicide tends to present high rates during their whole life cycle, which suggests that the persistence of violence is specific to each generation.

The region dummies show a positive and significant coefficient for all ages, indicating that Rio de Janeiro and São Paulo have significantly superior homicide rates to those of Minas Gerais for virtually all ages, even when controlled by the economic variables. The dummy coefficients suggest that this difference is more important for the younger age-groups, 15 to 17, after which it gradually loses importance with age. This result is interesting because the phenomenon of violence in the state of Minas Gerais is still not as disseminated as in the other states, and thus, the data seems to suggest, if our interpretation is correct, that the youth of Minas Gerais are still in the initial stages of entrance into illegal activity. The decision to join crime appears to be made at a slightly higher age.

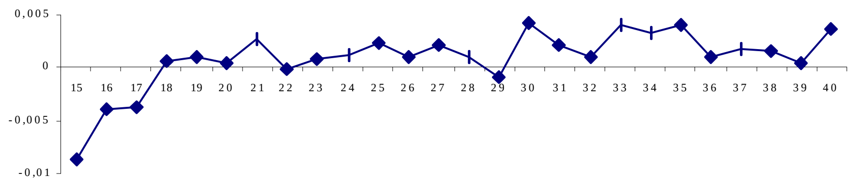
The significant reduction in the difference between Rio de Janeiro and São Paulo when data are controlled by economic variables and the inertia effect surprised us. For the non-controlled data, the difference fluctuates between about 100%, but for the controlled data this difference drops to about 20%. And this is spite of the fact that the data are not controlled by differences in public safety policies, which may present significant differences in both states during this period. Perhaps this difference is not so significant, or at least its impact may not be so. Perhaps, still, the public safety policy is linked to the economic variables: states that are wealthier during certain periods would have greater access to efficient safety tools. To reach an understanding of these possibilities, however, access to reliable data about public safety policies, which we did not have, is required.

With the insertion of the year variable as well as that of the price indexes, the coefficients (except for the constant) showed practically no alteration. The insertion of the Gini coefficient into the model does not notably alter the economic variable coefficients either. The coefficient that shows greatest change is the one for the São Paulo state dummy, an expected result, as these variables show a positive correlation. The Gini coefficient displays a positive sign and is significant for all ages, however, it is not possible to clearly distinguish the effects of unequal income in specific states. What can be observed in the data is evidence that the state of Rio de Janeiro is more violent than the state of São Paulo, and the two states in turn are more violent that the state of Minas Gerais. These differences, however, may not be creditable solely to income inequality. In order to study this effect in a more controlled manner, it would be necessary to expand the number of regions examined so as to increase the variability of income inequality.

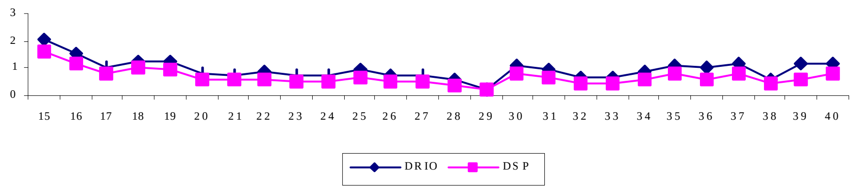
Graph 8: Unemployment Coefficient according to Age
Model 7



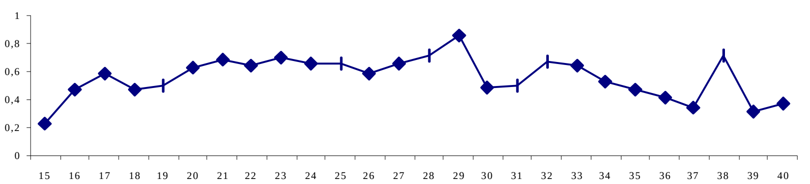
Graph 9: Real WageCoefficient according to Age
Model 7



Graph 10: Dummies Coeffcients according to Age
Model 7



Graph 11: Lagged Probability Coefficient According to Age
Model 7



V. Relationship with the existing literature on the economy of crime

There is a vast literature which seeks to explore the relationship between economic variables and the crime rate in the United States. In the Brazilian literature, however, most of these works concentrate on the areas of criminology and its social aspects. Perhaps this characteristic of the literature in Brazil is due partly to the scarce availability of a reliable data base, especially concerning public safety policies.

The greater part of the American works on economy are based on the theory of incentives initially proposed by Becker (1968), in which criminal activity is considered a substitute of legal activity.[22] In this model, individuals make decisions at every moment about allocating their time to legal activity or allocating it to criminal activity. The choice of which activity to engage in depends on the expected net gains from each activity. When calculating the expected net gains from crime, individuals consider their material gains and losses as well as the likelihood of arrest and punishment for criminal activity, and also the long term costs to their reputation and access to the labor market in case they are sentenced.

Recently, some authors have tried to incorporate the idea of *inertia* or *persistence* of the violence cycles. In the model constructed by Sah (1991), the probability of being punished for a criminal act is endogenous, and results in crime rate persistence. In the traditional groundwork, this probability is exogenous, constant along time and identical between individuals. As individuals decide whether or not to participate in crime, they consider their individual perception of the probability of punishment, and not the real probability. The individuals prior about the probability of punishment depends on the behavior of people that these individuals can observe during each period, while the real probability depends on the level of spending with public safety and the rate of participation in crime. Each period, a new cohort becomes a part of the total number of people taking

[22] Tauchen, Helen and Witte Ann D. (1994) and Grogger (1997), use a slightly different foundation. In the proposed model, criminal activity is not a substitute of legal activity. Individuals may engage in both types of activities simultaneously. In these kinds of models the choice of the individual is not between becoming involved or not in crime, but which is the optimal time to allocate to this activity.

part in the decision process of whether to participate or not in crime. The crime participation rate in period **T** depends on all active cohorts. In this sense, as agents get older, the body of information grows and the prior of the individuals tends to get closer to the real probability.

Sah assumes that for a given level of spending by the system there is a maximum level of arrests by the police, so that an increase in criminality reduces the real probability of punishment, and also influences the prior of individuals in relation to future choices. Thus, if the crime rate in the past was high, the tendency of an individual to join crime in the present will be greater, generating a violence persistence effect. The effect of past crime rates on present criminality depends on the active cycle of each cohort.[23]

There are basically three kinds of empirical approaches towards an understanding of the relationship between economic incentives and crime rate: the time-series analysis, the cross-section analysis and research based on individual behavior. The results found in these three types of approach differ largely. Freeman (1994) summarizes the main results for each of the three approaches.

In the case of the time-series analysis, the results depend on the period in which they were estimated, on the region and on how crime rate was measured. The results do not present a sole rule for the relationship between economy and criminal activity. However, this appears to be the data base that most adequately explains how, in a given geographic region, crime rate responds to economic fluctuations.

As for cross section analyses, they commonly use data from several locations this approach is better suited for explaining the crime rate variance, rather than its determinants. In this type of approach, the results are rather robust and crime rate presents an anti-cyclical behavior. The relationship between income inequality and crime also seems robust in this kind of analysis. Fanzylber, Lederman and Loayza (1999), using a data base from 45

---

[23] Another argument found in existing literature that may be used to explain the persistence of the violence cycles is the increase in social interaction proposed by Glaeser, Sacerdote e Scheinkman (1996). These authors were interested in explaining the high variance of crime rates along time and between regions, not explained by the variance in economic conditions. According to these authors, there is a positive co-variance between the decisions of agents to participate in crime and thus crime variance is a multiple of variance, provided the agents make independent decisions.

countries, found a positive relationship between income inequality , measured by the Gini coefficient, and the rate of intentional homicides. The weakness of this sort of analysis is that the results may be, to some extent, associated to different non-controlled characteristics of the populations in the experiments, thus generating a spurious correlation between the variables.

The individual data models show the strongest relationship between economic variables and the decision to participate in crime. This type of research, usually conducted in penitentiaries, does not consider the totality of the individuals but restricts itself to those that have already opted for criminal activity.

The abundance of such works reflects the gravity of the criminality problem as well as the problems in dealing with the subject from an economic perspective. Among the main obstacles are: 1) the lack of such a measure of gains from crime that allows the separation of the effects of less legal opportunities from the effects of high likelihood of punishment and increase in income from criminal activity; 2) the lack of exact measures of crime rate given the high number of sub-records; 3)problems measuring punishment probability, since the variables commonly used such as police force expenditures, number of police officers per capita and rates of arrest, may be correlated to the increase in violence or even present measuring mistakes. One stylized fact in the existing literature is a link between higher gains from crime and economies where there is a greater production and commercialization of drugs. Some authors defend that crime has increased in the last two decades due to the proliferation of drug use.

The results of three recent works on the American economy are in accord with the evidence found in this article. Grugger (1997) demonstrates that the behavior of young males shows a high degree of response to economic incentives and that therefore the decrease in real wage in the last two decades is an important component in the crime rate elevation. Tauchen and Witte (1994), on the other hand, demonstrate that those youths who allocate a larger amount of their time to work or school show a reduced probability of participating in crime. Both investigations used a seven-year period data base concerning a cohort of males born in 1945 in Pennsylvania. Freeman (1996) argues that the crime rate has grown due to the reduction of labor opportunities for young males of inferior qualifications. According to the author, the gains from legal activity for less qualified individuals were reduced in contrast to the increased gains from crime. Aside from this, the labor supply elasticity in response to wage variations is significant enough to increase the tendency to criminal activity. Add to that the argument that criminal activity may not necessarily be a substitute for legal activity, but may be conducted in parallel, in order to augment income.

As for the Brazilian case, results are still uncertain. Carneiro and Phebo (1999) find results that are rather different from the expected. The authors analyze the relationship between crime rate - which they measure as homicides per 100,000 inhabitants in the municipalities of Rio de Janeiro and São Paulo - and economic fluctuations, through a unbalanced panel data model for four years and estimated by weighed least squares. The results indicate a negative relationship between income inequality and crime rate, as well as a pro-cyclical behavior of criminality.

Beato and Reis (1999) analyze the correlation between crime rate, which they measure separately as crimes against property and crimes against individuals, and economic development in the municipalities of the metropolitan region of Belo Horizonte in a cross section model.[24] The correlation found depends on the way in which the crime rate is measured. This work, however, did not seek to control all independent variables simultaneously, but only examined the correlation and dispersion between each pair of variables.

A likely problem these two works have in common is the units of analysis. The integration of the labor market in the metropolitan regions probably prevents the economic fluctuations from being differentiated on the municipal level, which may result in a spurious correlation, as discussed in the introduction. Let's say, for instance, that two neighboring regions display significantly different crime rates. The municipality with the greater crime rate, in the absence of migration, should show some kind of compensation in

[24] As proxies for level of development, the authors use the rates of human development, illiteracy and child mortality.

terms of well-being for its inhabitants, for instance a higher wage or a lower unemployment rate. Thus, a cross section data base which indicated the existence of a positive correlation between, for example, real wages and crime rate, does not mean that an increased salary increases the crime rate. In fact, it might just mean that the opposite is true: in case the real wages increase, criminality might drop without contradicting the correlation obtained in the cross-section. The result obtained may mean nothing but that if one city presents a greater crime rate, the life quality of the inhabitants is compensated by an increase in real wages. For this reason, it seems important to consider a time-series data base, in which simultaneous fluctuations of economic variables during different periods of time in all regions are observed.

Besides this, as discussed in the introduction, the use of homicide rate per 100,000 inhabitants as a data base may skew the results due both to the existence of distinct sex and age compositions in the different regions and to the omission of the age variable. For this reason, in this work we constructed the data base from data calculating the frequency of death by homicide for each age, year and region.

The construction of this data base allowed us to consider two aspects which appear extremely relevant in the understanding of the violence cycles: incorporating the inertia effect and estimating specific coefficients for each age. When incorporating the inertia effect as specific to a given generation, the inertia of the violence cycles, previously observed in other works, acquires a new interpretation. The most robust result found in this article refers precisely to the importance of the lagged probability to explain the violence cycles, and was obtained through the arranging of the homicide death rates by cohorts. The estimated coefficients for lagged probability suggest that the individuals decide to enter crime between 15 to 20 years of age, and that once they decide to participate in this activity, they tend to continue in it. It is outside the scope of this work to explain the determinants of this inertia.

The estimate of the model for each age group allows us to suggest a some relationships between economic cycles and cycles of violence. The results evidence that the economic variables are relevant to explain criminality until a maximum of 20 years of age.

This result conforms to the works of Gruegger and Tauchen and Witte, and seems of great importance from a public intervention point of view. The facts suggest that policies aimed at the young male population may reduce criminality in the long term. However, since the impact of these policies is persistent along the life cycle their effect can only be perceived in the long term, which may render their implementation less attractive to politicians in office.

The results found for the relationship between income inequality and violence, though they display a positive and significant sign, are not very robust. The enormous disparity observed in the violence levels of the three states may be partly explained by unequal income levels, but that result may be associated to state-specific characteristics related to public safety policies. As we have emphasized before, however, the lack of such a reliable data base made it impossible to extend this work into a further investigation of that relationship.

VI. Research Schedule

Though the results encountered in this work seem robust, the precarious economic data for each cohort, and the lack of data about public safety policies, undermine the tested models. It is possible to point out at least three extensions of this work:

1)The construction of a data base where economic variables reflect the characteristics of each age group. In this article, the economic variables were approximated using the average;

2)Inclusion of instrumental variables for public safety policies;

3)Expansion of the number of federal units analyzed so as to obtain a greater dispersion and income inequality and test alternative inequality measurements.

In principle, the relevant variables for the ingress into illegal activities should be contingent upon individual characteristics, and not upon the average of each region's population for each year. An interesting rearrangement would be to make the homicide rate data contingent upon level of education and parental characteristics, and to make economic variables contingent upon age and level of education. Unfortunately, the data available by the SIM show many blanks precisely of this type of data. An intermediary solution would be to make the economic variables, at least, contingent upon age, besides region and year.

As far as the public safety policies, some alternative variables such as the number of police officers might be attempted as proxies for the government commitment to punish crime. Public safety policies, however, are not comprised solely of punishment, and it isn't even clear to what degree the number of officers adequately represents this commitment. This aspect needs a more careful investigation of criteria and better indicators of public safety policies.

Finally, in reference to the relation between income inequality and death homicide rates, two extensions seem appropriate. On one hand, expansion of the number of federal units in order to distinguish the income inequality from other regional aspects and, on the other, testing of alternative income inequality indexes. The main difficulty with this extension lies in the high number of non-registered death in some Brazilian states, especially in the north and northeast.

VI. Referências Bibliográficas

Amemya, T. (1985): *Advanced Econometrics*; Oxford, Basil Blackwell.

Andrade, M. V. e M. B. Lisboa (2000): "Alguns Fatos Estilizados sobre o setor saúde no Brasil" (mimeo)**,** EPGE/FGV.

Beato, C. e I. A. Reis (1999): "Desigualdade, Desenvolvimento Sócio Econômico e Crime"; Mimeo. Texto apresentado no seminário *Desigualdade e Pobreza no Brasil*, promovido pelo IPEA/RJ, Rio de Janeiro, agosto, 1999.

Becker, G. (1968): "Crime and Punishment: An Economic Approach"; *Journal of Political Economy*, **76(2):**169-217.

Carneiro, L.P. e L. Phebo (1999): "Magnitude, Custos econômicos e Políticas de Controle da Violência no Rio de Janeiro", Banco Interamericano de Desenvolvimento. Série *Documentos de Trabalho R347.*

Fajnzylber, P.; D. Lederman e N. Loayza (1999): "Inequality and Violent Crime". ???

Fajnzylber, P; D. Lederman e N. Loayza (199?): "Determinants of Crime rates in Latin America and the World" *World Bank, Latin American and Carribean Studies, Viewpoint Series.*

Freeman, R. B. (1994): "Crime and the Job Market"; National Bureau of Economic Research, *Working Paper 4910*. Cambridge, March.

Freeman, R. B. (1996): "Why do so many young American men commit crimes and what might we do about it?" National Bureau of Economic Research, *Working Paper 5451*. Cambridge, February.

Fundação Seade (1993): Anuário Estatístico do Estado de São Paulo.

Governo do Estado do Rio de Janeiro (1997): Anuário Estatístico do Estado do Rio de Janeiro**;** Secretaria de Estado e Planejamento.

Glaeser, E.; B. Sacerdote e J. Scheinkman (1996): "Crime and Social Interactions"; *Quartely Journal of Economics*, **111(2).**

Grogger, J. (1997): "Market Wages and Youth Crime"; National Bureau of Economic Reseach, *Working Paper 5893*. Cambridge, March.

Instituto Brasileiro de Geografia e Estatística: Censo Demográfico, 1980 e 1991.

Instituto Brasileiro de Geografia e Estatística (1996): Contagem Populacional; Cd-rom.

Li, M. (1977): "A Logit Model of Homeownership"; *Econometrica*, **45(5).**

Maddala, G.S. (1983): *Limited Dependent and Qualitative Variables Econometrics*; New York, Cambridge University Press.

Ministério da Saúde (1997): *Sistema de Informação sobre Mortalidade 1979 – 1997. Dados de Declaração de óbito*; CD-ROM.

Ministério da Previdência e Assistência Social (1990): *Nona Classificação Internacional de Doenças.*

Ryder, N. B. (1965): “The Cohort as a concept in the study of Social Change”; *American Sociological Review*, December, **30(6).**

Sah, R. K. (1991): “Social Osmosis and Patterns of Crime”; *Journal of Political Economy*, **99(6).**

Tauchen, H. e A. Witte (1994): “Work and Crime: An Exploration using Panel Data”; National Bureau of Economic Research, *Working Paper 4794,* Cambridge.

Shryock, A.S. e Siegel, J.S. (1973): *The Methods and Materials of Demography*; U.S. Bureau of Census. Second Printing, Washington D.C.

Apendix1: Regression Results: Models 7 and 10

Tabela 7: Modelo 7

Idade	Cte c(1)	Salreal c(2)		Des c(3)		Drio c(4)	Dsp c(5)		Prob(-1) c(6)	
15	-5.68	-0.0088	(**)	-0.181	(**)	2.04	1.59	(**)	0.23	(**)
16	-3.76	-0.0039	(**)	-0.110	(**)	1.52	1.20	(**)	0.47	(**)
17	-2.65	-0.0037	(**)	-0.079	(**)	1.06	0.77	(**)	0.58	(**)
18	-4.19	0.0007	(**)	-0.053	(**)	1.26	1.01	(**)	0.48	(**)
19	-4.17	0.0011	(**)	-0.036	(**)	1.25	0.98	(**)	0.50	(**)
20	-2.84	0.0005	(**)	-0.025	(*)	0.80	0.58	(**)	0.63	(**)
21	-2.89	0.0027	(**)	-0.006	(***)	0.76	0.59	(**)	0.69	(**)
22	-2.98	-0.0001	(**)	0.007	(***)	0.87	0.56	(**)	0.64	(**)
23	-2.50	0.0008	(**)	-0.007	(***)	0.70	0.53	(**)	0.70	(**)
24	-2.89	0.0012	(**)	-0.003	(***)	0.70	0.49	(**)	0.65	(**)
25	-3.16	0.0023	(**)	0.010	(***)	0.95	0.68	(**)	0.66	(**)
26	-3.39	0.0010	(**)	-0.019	(***)	0.73	0.50	(**)	0.58	(**)
27	-3.19	0.0022	(**)	0.026	(*)	0.71	0.48	(**)	0.65	(**)
28	-2.52	0.0010	(**)	0.108	(***)	0.58	0.38	(**)	0.71	(**)
29	-0.90	-0.0008	(***)	-0.021	(***)	0.24	0.19	(***)	0.86	(**)
30	-4.63	0.0042	(**)	-0.008	(***)	1.11	0.80	(**)	0.49	(**)
31	-4.38	0.0021	(**)	-0.003	(***)	0.93	0.67	(**)	0.50	(**)
32	-2.89	0.0010	(**)	-0.006	(***)	0.69	0.43	(**)	0.67	(**)
33	-2.95	0.0040	(**)	-0.005	(***)	0.69	0.43	(**)	0.64	(**)
34	-4.36	0.0032	(**)	0.018	(***)	0.86	0.58	(**)	0.52	(**)
35	-4.83	0.0041	(**)	0.003	(***)	1.08	0.81	(**)	0.47	(**)
36	-4.90	0.0009	(***)	-0.020	(***)	1.01	0.57	(**)	0.41	(**)
37	-5.60	0.0018	(***)	-0.034	(***)	1.19	0.82	(**)	0.34	(**)
38	-2.70	0.0016	(***)	0.006	(***)	0.55	0.44	(**)	0.72	(**)
39	-5.69	0.0005	(***)	0.006	(***)	1.15	0.62	(**)	0.32	(**)
40	-5.79	0.0038	(**)	0.000	(***)	1.18	0.82	(**)	0.37	(*)

(*) Significant at 5%

(**) Significant at 1%

(***) Not significant

Table 2: Model 10

Age	Cte	wage		Unemployment		GINI		Drio	Dsp		Prob(-1)		INPC	
	c(1)	c(2)		c(3)		c(4)		c(5)	c(6)		c(7)		c(8)	
15	-7.43	-0.009	(**)	-0.179	(**)	2.20	(***)	2.18	1.83	(**)	0.19	(**)	0.0011	(***)
16	-6.07	-0.004	(**)	-0.094	(**)	3.44	(**)	1.59	1.44	(**)	0.45	(**)	0.0018	(***)
17	-3.96	-0.004	(**)	-0.079	(**)	2.15	(*)	1.11	0.93	(**)	0.57	(**)	-0.0017	(***)
18	-6.95	-0.001	(***)	-0.053	(**)	4.00	(**)	1.45	1.37	(**)	0.42	(**)	-0.0008	(***)
19	-7.12	-0.001	(***)	-0.038	(**)	4.81	(**)	1.39	1.33	(**)	0.46	(**)	-0.0033	(*)
20	-5.30	0.000	(***)	-0.018	(***)	3.32	(**)	0.98	0.88	(**)	0.57	(**)	0.0006	(***)
21	-4.84	0.002	(*)	-0.004	(***)	2.56	(**)	0.92	0.84	(**)	0.63	(**)	0.0003	(***)
22	-6.05	-0.001	(***)	0.010	(***)	4.45	(**)	1.06	0.93	(**)	0.58	(**)	-0.0019	(***)
23	-4.70	0.000	(**)	0.000	(***)	2.97	(**)	0.85	0.78	(**)	0.65	(**)	0.0004	(***)
24	-6.13	0.000	(***)	0.008	(***)	4.42	(**)	0.92	0.87	(**)	0.57	(**)	0.0000	(***)
25	-7.40	0.001	(***)	0.025	(*)	5.54	(**)	1.26	1.18	(**)	0.54	(**)	0.0005	(***)
26	-5.03	0.000	(***)	-0.019	(***)	2.23	(*)	0.87	0.71	(**)	0.53	(**)	-0.0019	(***)
27	-6.11	0.002	(*)	0.044	(**)	3.90	(**)	0.87	0.79	(**)	0.59	(**)	0.0022	(***)
28	-3.61	0.001	(***)	0.016	(***)	1.24	(***)	0.67	0.50	(**)	0.67	(**)	0.0008	(***)
29	-2.18	-0.001	(***)	-0.019	(***)	1.65	(***)	0.33	0.33	(*)	0.82	(**)	-0.0004	(***)
30	-8.54	0.003	(**)	0.011	(***)	5.42	(**)	1.32	1.20	(**)	0.41	(**)	0.0012	(***)
31	-6.74	0.002	(***)	0.006	(***)	3.02	(**)	1.10	0.93	(**)	0.43	(**)	0.0005	(***)
32	-4.91	0.001	(***)	0.007	(***)	2.16	(*)	0.85	0.65	(**)	0.59	(**)	0.0026	(***)
33	-5.53	0.001	(***)	0.010	(***)	2.89	(**)	0.89	0.70	(**)	0.56	(**)	0.0037	(***)
34	-6.53	0.004	(**)	0.038	(***)	2.58	(*)	0.98	0.80	(**)	0.48	(**)	0.0049	(**)
35	-7.21	0.005	(**)	0.023	(***)	2.69	(**)	1.23	1.04	(**)	0.40	(**)	0.0045	(*)
36	-7.04	0.001	(***)	-0.012	(***)	2.78	(*)	1.15	0.80	(**)	0.35	(**)	0.0007	(***)
37	-7.71	0.002	(***)	-0.015	(***)	2.08	(***)	1.36	1.02	(**)	0.26	(**)	0.0038	(***)
38	-4.43	0.002	(***)	0.013	(***)	1.64	(***)	0.70	0.62	(**)	0.64	(**)	0.0025	(***)
39	-8.24	0.001	(***)	0.017	(***)	4.01	(**)	1.18	0.86	(**)	0.31	(**)	0.0033	(***)
40	-6.82	0.003	(***)	0.000	(***)	1.49	(***)	1.24	0.93	(**)	0.34	(*)	-0.0008	(***)

(*) Significant at 5%
(**) Significant at 1%
(***) Not significative