Age Structure Explaining a Large Shift in Homicides: The Case of the State of São Paulo

VERY PRELIMINARY. PLEASE DO NOT CITE

FIRST VERSION: August-2007

João M P De Mello† and Alexandre Schneider‡

Abstract

After reaching a historic peak by the end of the 1990s, homicides in large cities in the state of São Paulo dropped sharply. Several explanations have been advanced, most prominently improvements in policing, adoption of policies such as dry laws, and increased incarceration. In this paper, we show that demographic changes play a large role in explaining the dynamics of homicide. More specifically, we present evidence of a strong co-movement between the proportion of males on the 15-25 age bracket and homicides at the statewide and at city levels, and argue that the relationship is causal. We estimate that a 1% increase in the proportion of 15-to-24-year-old males causes a 4.5% increase in homicides.

KEYWORDS: Age Structure, Demographic Change, Homicides

† Departamento de Economia, PUC-Rio. Corresponding author: jmpm@econ.puc-rio.br.

‡ Secretário de Educação do Município de São Paulo. Mr. Schneider would like to stress that opinions expressed here are solely from the authors, and do not reflect any official position of the Mayorship of São Paulo.

I. Introduction

After increasing steadily over the 1990s, homicides in the state of São Paulo fell sharply in the 2000s. As an illustration, there were 24 homicides per 100thd inhabitants over the year 2005 at the São Paulo Metropolitan Area (SPMA), down from 52 in the 1999 peak, and more than 20% *less* than the level in 1992. The dynamics are similar to that of the large American cities, where homicides peaked ten years before. Better than words, consider the following four figures:

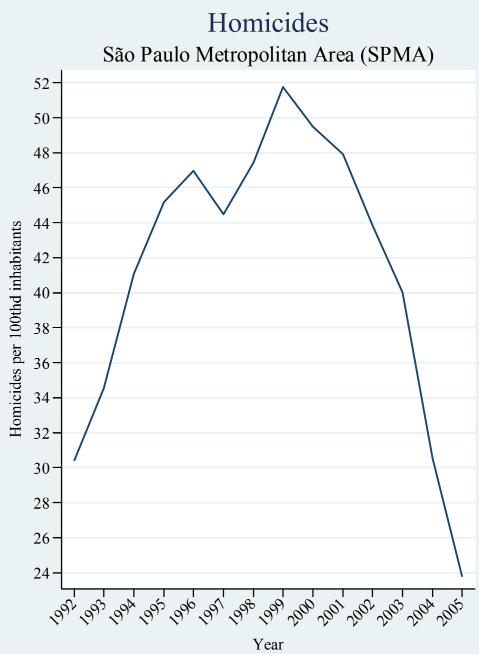


Figure I: Aggregating all 39 cities in the SPMA

Source: Secretaria de Segurança Pública de São Paulo (SSP-SP)

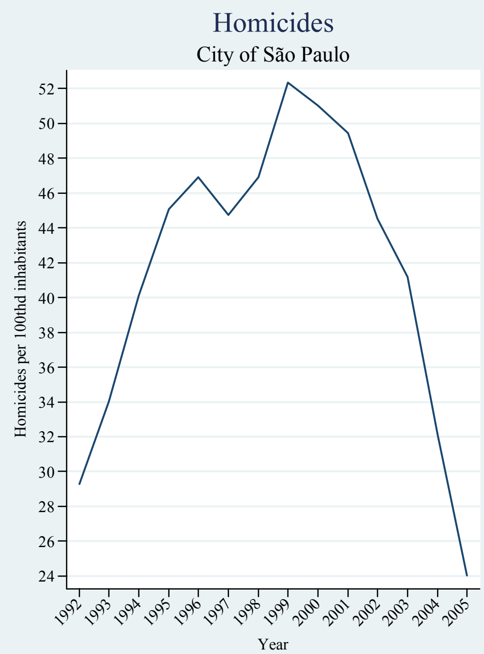


Figure II

Source: Secretaria de Segurança Pública de São Paulo (SSP-SP)

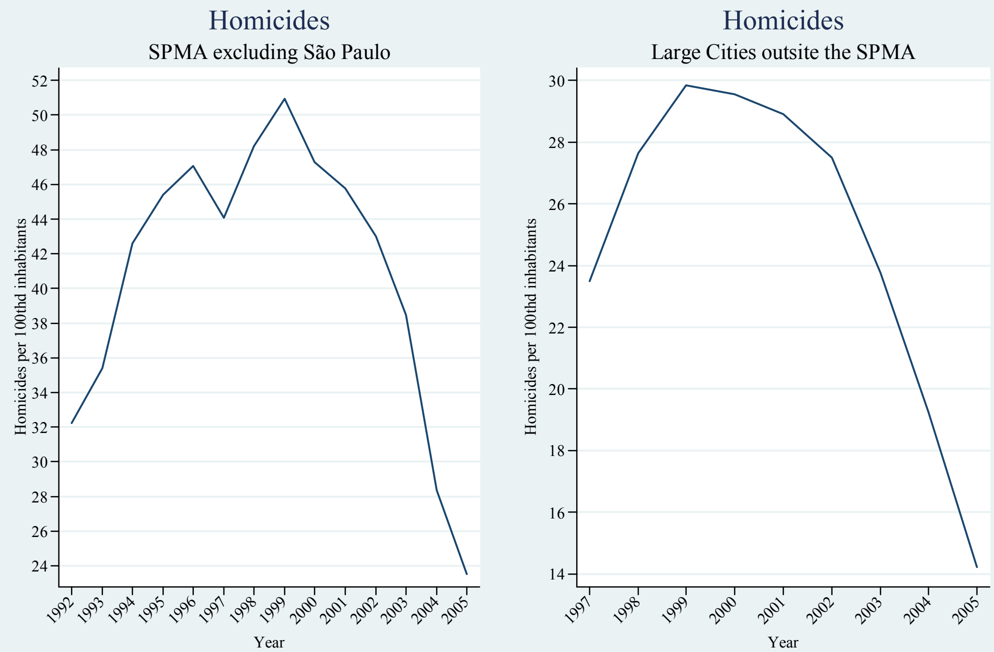


Figure III: Aggregating all 38 cities in the SPMA except São Paulo
Source: Secretaria de Segurança Pública de São Paulo (SSP-SP)

Figure IV: All Cities larger then 100thd inhabitants
Source: Secretaria de Segurança Pública de São Paulo (SSP-SP)

Figures I-IV summarize the motivation for this paper. Although the spotlight was always on the SPMA in general, and the city of São Paulo in particular, the homicides movements are similar across the board, at least in middle sized to large cities (more than 100thd inhabitants).[1] Comparing figure IV with the other three figures, one can see considerable cross-city variation in levels but, overtime, there is a remarkable similarity in the dynamics of homicides (see table I for more details): in all four groups depicted above homicide increased throughout the 1990s, peaked in 1999 and fell dramatically thereafter. Our research question is straightforward: what role did demographics play in explaining dynamics of homicide in the state of São Paulo over this period?

As usual, there are many other culprits. Not surprisingly, the government of the state of São Paulo claim credit.[2] Indeed, several improvements in policing occurred somewhat concurrently. Among them, the adoption of a unified data and intelligence system, INFOCRIM (a version of *Compustat*), and the implementation of a photo database of criminals are the most prominent. Incarcerated population and the number of

[1] Graphs for all cities with more then 100thd inhabitants are available upon request.
[2] As we shall see shortly, policing is done mainly at state-level in Brazil.

police officers also increased over the period although, as usual, it is hard to disentangle cause and effect. Among municipal-level policy measures, several others are worth mentioning: the adoption "dry laws" (which are restrictions on the recreational sales of alcohol), the creation of municipal police forces, the adoption of DISQUE-DENÚNCIA (an anonymous hotline to denounce crimes), and efforts to enforce the banning of possession of armed guns (see Kahn and Zanetic [2005]).[3]

As we will see, while it is conceivable that all these measures indeed helped reducing homicides, they cannot account for the dynamics observed in the data, neither one-by-one nor as a whole. "Dry laws", for example, can be dismissed on two grounds. First, their adoption is not nearly as widespread as the drop in homicides in the end of the 1990s. Second, adoption occurred *after* 2000. Since there is evidence that they caused a non-negligible reduction in homicides, it is likely that "dry laws" were a contributing factor, strengthening a trend that was already in place but it is simply impossible that they might have caused it.[4] In section IV, we provide a timeline of adoption of other policy measures and argue, one by one, that they cannot possibly explain the data.

We show that demography played a significant role in explaining the dynamics of homicide in the state of São Paulo. There is a remarkable co-movement between homicides and of the proportion of male between 15 and 25 years olds, the age-gender group most inclined to committing crimes.[5] Over the 1991-2000 period, when homicides rose 63.2%, the SPMA gained some 216,000 15-to-25-year-olds, a 15.3% increase against a population increase of 11.6. In contrast, over the 2000-2005 period, while its population increased 6.6%, the SPMA *lost* some 60,000 young males.[6]

---

[3] From 2003, the "Disarmament Act" (Lei do Desarmamento, Lei Nº 10.826) is a federal act that increased significantly the penal and civil costs of illegal possession and trade of fire guns. See http://www.camara.gov.br/internet/infdoc/Publicacoes/html/pdf/Desarmamento.pdf

[4] By comparing homicide dynamics in adopting and non-adopting municipalities in the SPMA, Biderman, De Mello and Schneider [2007] estimate that "dry laws" caused a reduction of roughly 12% in homicides.

[5] At the individual level, criminal involvement and age is one of the most robust relationships in all social sciences, dating back to at least Goring [1913]. A very non-exhaustive list of more recent work would include Wilson and Herrnstein [1985], Blumstein [1995] and Cook and Laub [1998].

[6] While crime has been increasing in the United States in general, large cities such as Chicago, Los Angeles and New York experienced a *drop* in crime. Among the several candidates, only improvements in policing and demographic change survive scrutiny. See Skogan [2003]. For example, the three cities combined lost some 200,000 young males over the 2000-2005 period. See also *The Economist*, "New Police Model," June 9th, 2007.

Evidently, this correlation is not conclusive. For one thing, the co-movement may be a sheer coincidence. When one analyzes particular cities, the same pattern arises in general, which is all the more suggestive but makes it more difficult to disentangle time effects that may be causing both homicides and demography from a clean causal effect. Fortunately, however, there is sufficient variation across cities on the time-series of demography to estimate the effect of demography on homicides. After controlling for year and city effects, we find a significant elasticity of homicides to the proportion of young males: a 1% increase in the proportion of males between 15 and 25 years old is associated with a 5% increase in homicides.

It is difficult to exaggerate the importance of this result. First, crime in general, and violent types such as homicides in particular, are a serious and costly social phenomenon. Second, there has been an enormous debate on the causes of the recent drop in homicides in the SPMA. Our results are negative in one sense and positive in another. Negative in the sense that policy played a secondary role. Positive in the sense of figure V: the spike in homicides observed in the 1990s is a side effect of the so-called "demographic bonus" that Brazil has been experiencing. Figures V.A, V.B and V.C show the distribution of males in the SPMA for three years: 1993, 1999 and 2005.

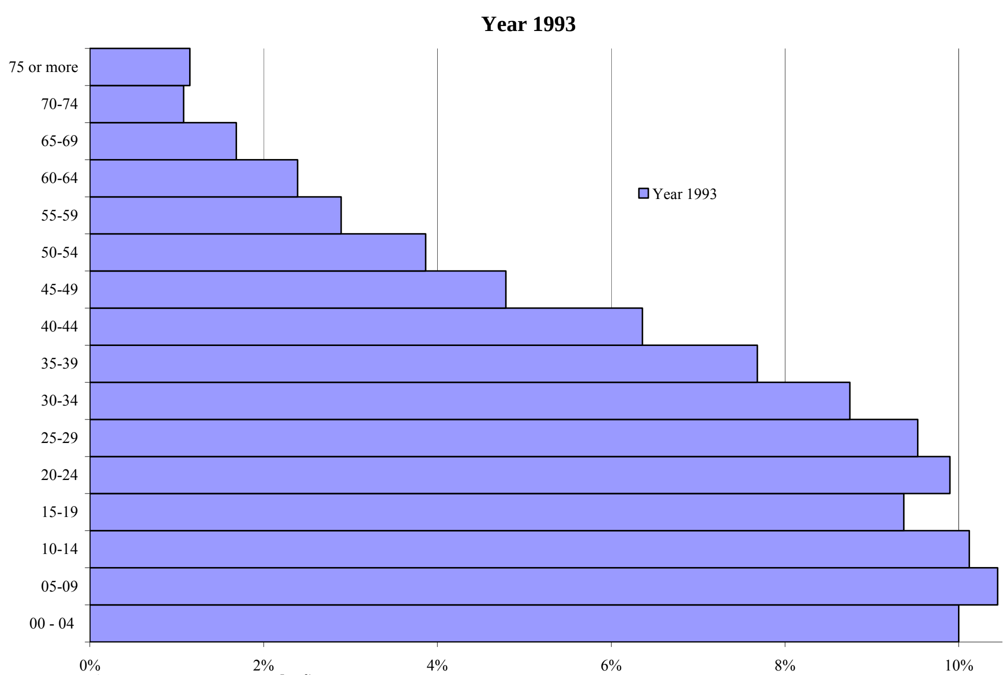


Figure V.A Source: Fundação SEADE

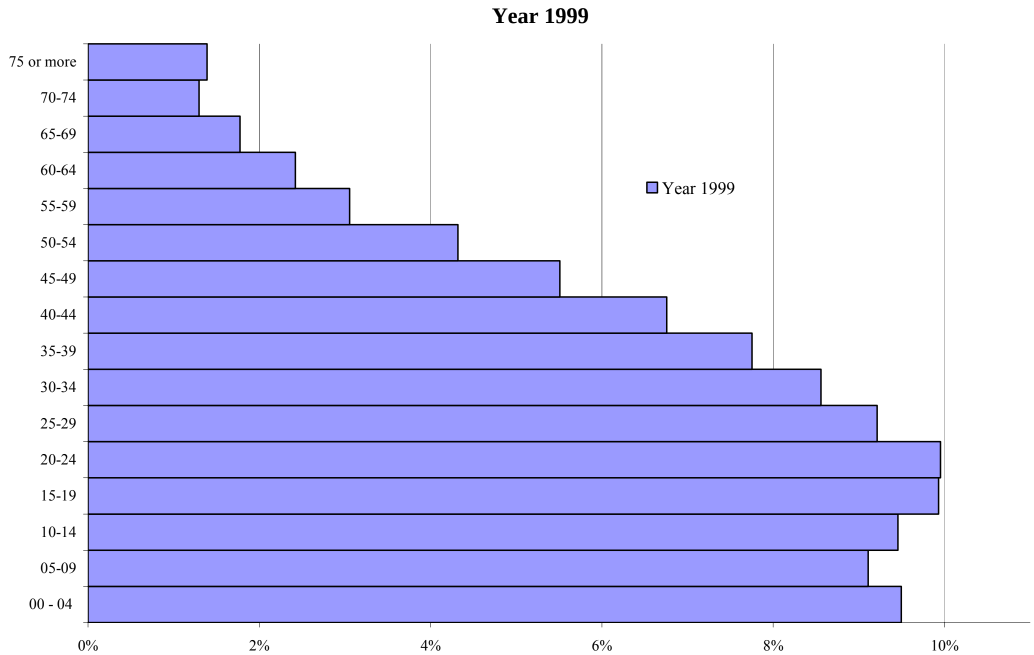


Figure V.B Source: Fundação SEADE

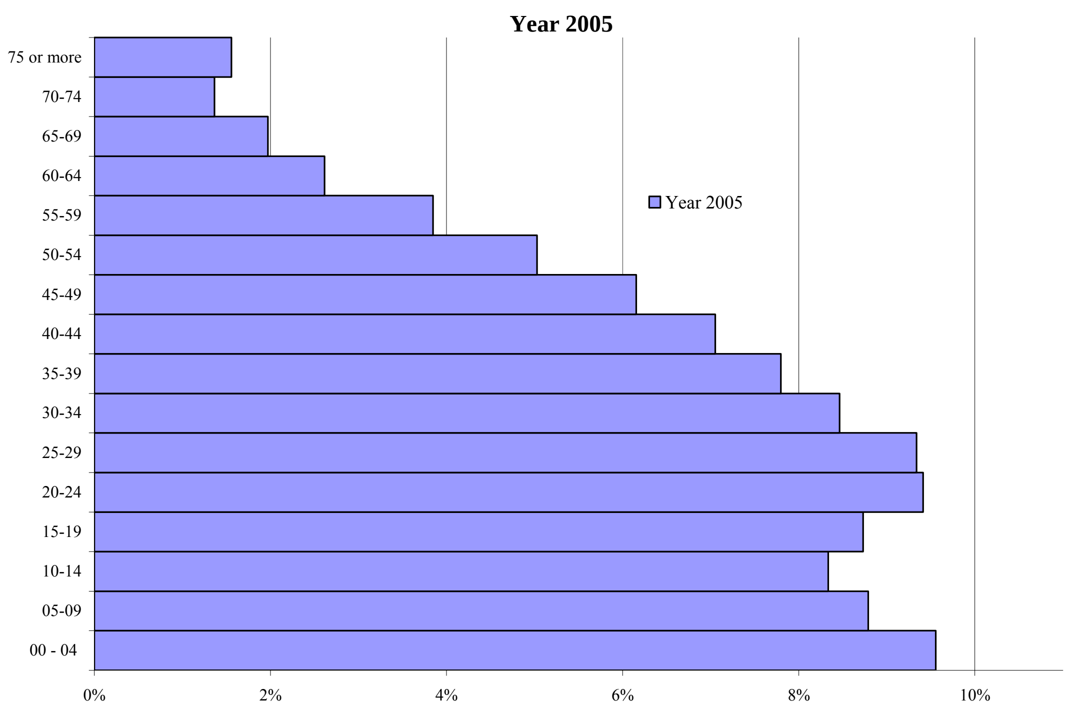


Figure V.C Source: Fundação SEADE

There is pronouncedly large cohort (reaching their thirties now) that resembles a "wave" moving up the demographic pyramid.

The case for demographics as an explaining factor demands that young males are particularly crime prone. While incarceration data is scant in São Paulo, there is a best alternative: hospital data, from which one can recover the age of homicide victims. Figure VI depicts the statewide distribution of homicide victims in the year 2000.[7]

Distribution of Homicide Victims in 2000, by Age and Gender

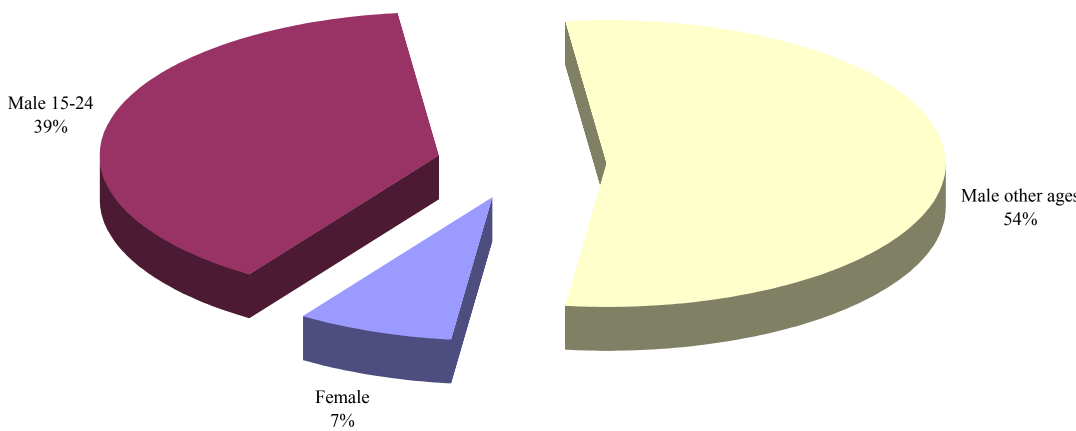


Figure VI Source: DATASUS

The paper is organized as follows. In section II, we describe the data used and show some preliminary results. Using victim data, we construct age-specific homicide rates, and perform a simple Oxaca-Blinder type decomposition. Keeping 1992 murder rates by age bracket constant, how much of the dynamics of homicides can be explained by changes in demographics? As we will see, although demographics fail to pin down the level, they explain some 70% of the time-series variation in homicides. When high-school drop-outs are factored in, demographics performance in explaining the level

[7] Ideally, one would like to observe the age of the perpetrator, not the victim. Nevertheless, perpetrators and victims are usually similar in socio-economic terms, allowing us to interpret figure VII as evidence that demographics, at a minimum, strongly correlate with homicides.

improves dramatically. This suggests that unsurprising fact that the effect of young male on crime is mitigated by having a viable non-criminal alternative.

Clearly, time may have caused both homicides and demography. For this reason, cross-city variation is added in section III. Regressions results show that the link from demographics to homicide survives the inclusion of time specific dummies. In fact, after controlling for time *and* city effects, a 1% increase in the proportion of young males (age 15 to 34) causes a ?? increase in homicide. Therefore, one can be confident that there is indeed a strong causal link from the proportion of male youths to homicide.

While our estimates are in line with long established the individual-level evidence from criminology, they contrast with some works that have failed to find a large impact of demography on aggregate crime rates with U.S. data (Levitt [1999]).[8] Although offenders are male and young, the proportion of crime prone individuals may not change sufficiently over time to influence aggregate crime rates.

Results can be somewhat reconciled by keeping in mind that, although being in a crime-prone age bracket matters, also important is being out of the school system. In this sense, our results are compatible with Levitt and Donohue [2001].[9] Or perhaps a changing age structure will matter most in poorer, extremely unequal, weak law-and-order environments such as the state of São Paulo.

In section IV we enumerate a long list of alternative explanations. We show that, while other factors may have contributed to the decline in murder rates in the 2000s, they cannot explain time-series pattern of the data, particularly the inversion that occurred in late 1990s. Section V concludes.

II. Data and the Time Series Decomposition

We use four sources of data. City-level homicide data comes from the Secretaria de Segurança Pública de São Paulo (SSP-SP), the state-level enforcement authority.[10]

[8] See footnote 5.
[9] Levitt and Donohue [2001] investigate the role played by legalized abortion in explaining the 1990 reversion of crime trends in the U.S., finding a large impact. Their (rather controversial) interpretation is that legalized abortion reduced the supply of crime prone individuals.
[10] In Brazil, by constitutional law, police enforcement is a state attribution.

City-level murder rates from 1991 to 2005 are available for the cities in São Paulo Metropolitan (SPMA), and from 1997 to 2005 for the rest of the large cities in the state.

As an alternative, we use state-level aggregate victim data, by age bracket, for the 1984-2004, from the Serviço Único de Saúde (SUS) database (DATASUS). The SUS is the federal government run national health system, subordinated to the Ministry of Health. DATASUS data is at the hospital level, which means that homicides are allocated to the city where the hospital in which the victim died is, which is not necessarily where the murder occurred. When using aggregate data over many cities, SSP-SP and DATASUS data are equivalent.[11] However, for the procedures using cross-city variation, one should be cautious with DATASUS data. Therefore, for all city-level procedures, SSP-SP data is used.

Demographic data are from the 1990 and the 2000 census, and the intermediate population, available from the Instituto Brasileiro de Geografia e Estatistica (IBGE).[12] Finally, city-level high school drop-out rate is from the Secretaria Estadual de Educação, the state-level authority.

Table I shows some descriptive statistics for SPMA and for all other cities with more 100thd inhabitant, for three periods: 1991-1995, 1996-2000, and 2001-2005. Two key variables are depicted: homicide rates and % of male population in the 15-24 age group. To keep consistency with the decomposition below, we use DATASUS data, and only male homicide victims are considered.[13]

[11] They are not precisely equal because some homicides do not go on hospitals´ records, but the figures are very similar.
[12] The IBGE is an autarky below the Ministry of Planning, equivalent of a Bureau of Statistics.
[13] Roughly 93% of all homicide victims are male, which makes the distinction irrelevant.

TABLE I: Means for three different periods

	SPMA †		*OTHER LARGE CITIES ‡*	
	Homicide Rate§	**% Male 15-24**	**Homicide Rate§**	**% Male 15-24**
1991-1995	41.79	9.37	15.83	9.36
1996-2000	54.66	9.58	25.54	9.58
2001-2005	45.95	9.18	23.87	9.37

†: Cities in the São Paulo Metropolitan Area

‡: Cities with more than 100thd inhabitants in 2000

§: per 100thd inhabitants

Source: DATASUS and Instituto Brasiliero de Geografia e Economia (IBGE)

In an unpolished way, table I contains the story of the paper. Homicides rates and the percentage of males between 15 and 24 years old move together over time. Suggestively, in large cities outside the SPMA, where the demography varied less, crime also varied less.

To assess how much the demography explains of the time-series pattern of crime we perform two Oxaca-Blinder-type decompositions. The DATASUS data has information on the victims' age, which can be used as proxy for age group specific homicide rates.[14]

Let H_{ta} be the homicide victim rate of age group a in year t. Victims are divided into nine groups, and t = 1991, 1992,…., 2004.[15] Let P_{at} be the proportion of the population that, at year t, is the age group a. The overall homicide rate at time t is approximately[16]:

$$H_t = \sum_a P_{ta} H_{ta}$$

Define H_{ta_τ} as the homicide rate that would prevail in year t if age group specific homicides are kept constant but the demographic distribution changed to that of year τ. In other words:

$$H_{ta_\tau} = \sum_a P_{\tau a} H_{ta}$$

[14] Levitt [1999] uses incarceration data, which is generated on the offender side, a better but unavailable measure.

[15] Groups are 0-4, 5-14, 15-24, 25-34,35-44,45-54,55-64,65-74 and over 75.

[16] The "actual" homicide is rate is the sum of homicides over the sum of population.

H_{ta_τ} is used a rough prediction of what would the homicide rate be in year τ if the demography changed but the level of homicides of year t were kept constant. Figures VIII and IX show predicted and actual homicides for the 1984-2004 and 1991-2004 periods, holding homicide victim rates of 1984 and 1991, respectively.

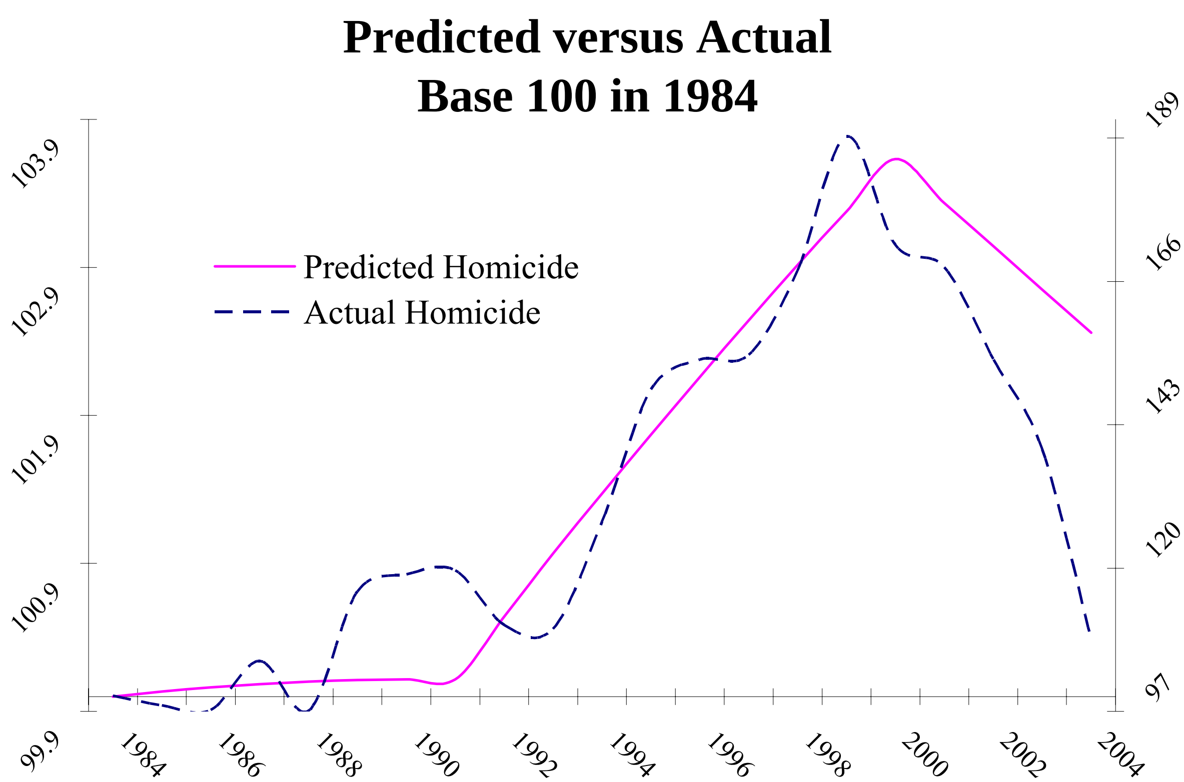


Figure VII Source: DATASUS and Instituto Brasileiro de Geografia e Estatística.

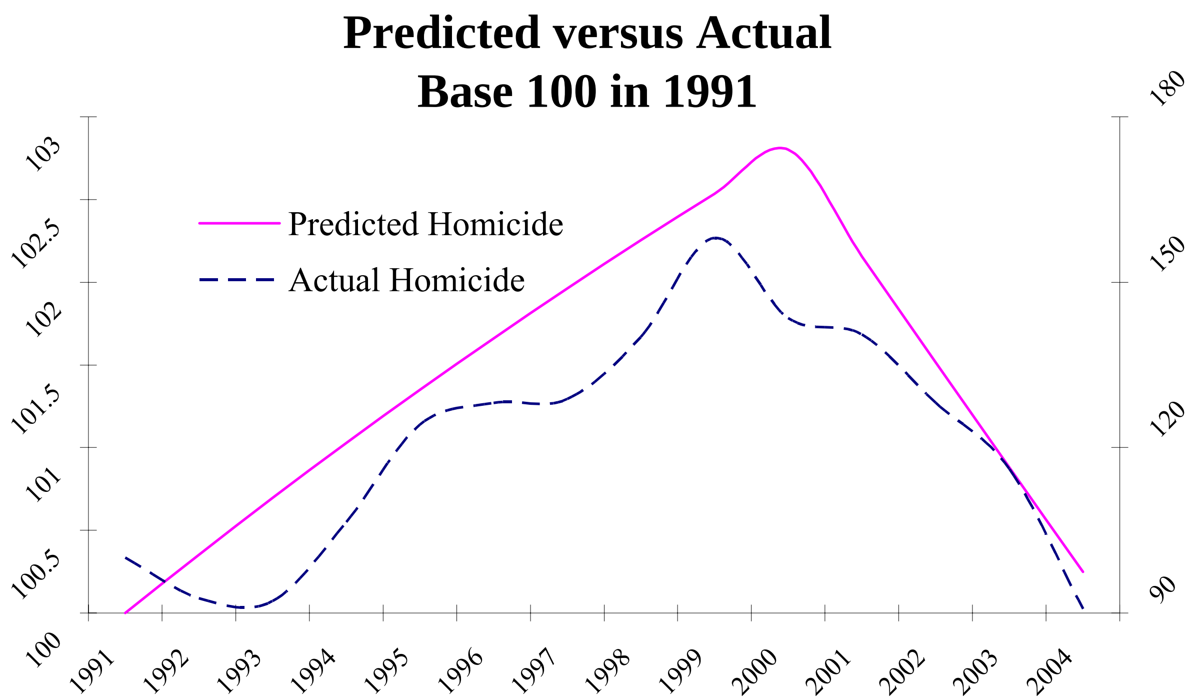


Figure VIII Source: DATASUS and Instituto Brasileiro de Geografia e Estatística.

Visual inspection of figures VIII and IX show two facts. First, there is a remarkable co-movement between the two series. This is true for the whole 1984-2004 period, and for the sub-period 1991-2004. Second, demographic changes do not pin down the level of homicides. Table II shows some OLS regressions of actual on predicted homicides. They serve two purposes. First, get a sense of variation of actual homicides are explained by predict homicides. Second, verify if their relationship co-integrates.

TABLE II: Actual versus Predicted Homicides
Dependent Variable: Actual Homicides, 1984-2004

	(1)	(2)	(3)	(4)†	(5)§
Predicted Homicide	20.07		24.29	6.76	22.46
	(2.53)***		(4.24)***	(2.56)**	(9.94)**
Year		780.95	813.78		
		(398.75)*	(168.76)***		
Year ^{2}		-0.20	-0.20		
		(0.10)*	(0.043)***		
Constant	-1306.98	-780447.30	-812341.90	-470.59	-0.90
	(176.14)***	(397307.10)*	(168172.80)***	(178.64)**	(1.96)
# Obs	21	21	21	21	20
R^2	0.73	0.58	0.88	0.20	0.23
Phillips-Perron‡	0.51			0.55	0.00

***: significant at the 1% level
**: significant at the 5% level
*: significant at the 10% level
†: Dependent Variable is the residuals of a fractional polynomial regression of actual homicides on four-degree polynomio of year
‡: Phillips-Perron *p*- value test for stationary of the residuals
§: First-Difference Regression
Source: DATASUS, Instituto Brasiliero de Geografia e Economia (IBGE) and authors' computations

In column (1) the raw relationship between predicted and actual is shown. As expected, predicted homicide explain a significant part of the variation in actual homicides: 73%. There is no evidence, however, that their relationship is stable: residuals are not stationary, which indicates that the two series do not co-integrate. The very large estimated coefficient reflects the fact that predicted homicides do not pin down the level of actual homicides. Visual inspection of figures VIII and IX suggests regressing actual homicides on year and year^2 (results in column (2)). Interestingly, predicted homicides do a better job at explaining actual homicides than the 2-degree polynomial of time (58% versus 78%). In column (3), instead of horseracing, the polynomial and predicted homicides are included. Three facts are noteworthy: 1) predicted homicides survive the inclusion of a polynomial of time; 2) predicted homicides explain, above and beyond the polynomial, an additional 20% of the variation in actual homicides; 3) the residuals are now stationary.

In column (4), a fractional polynomial regression in implemented: actual homicides are regressed against a four-degree polynomial.[17] Residuals were then regressed against predicted homicides. Predicted homicides are still related to actual homicides and explain some 20% of its variation, a number similar to the one in column

[17] The procedure chooses an optimal polynomial of four terms. Details are available upon request.

(3). The two series, however, are not co-integrated (Phillips-Perron test for residual stationary p – value = 55%). Finally, in column (5) we run the model in first differences because both actual and predicted are integrated of order 1. The coefficient on predicted homicides is similar to those in columns (1) and (3) and, when series are differentiated they co-integrate.

Evidently, this should not be viewed as "causal" effect of demographic changes. For one thing, these models are statistical models, with little economic content. Second, there is no reason to believe that changing demographic will not affect the level of homicide itself. In fact, social interaction effects already documented in the literature suggest that, an increase in the crime prone age population should increase the level of homicide rates (see Glaeser, Sacerdote and Scheinkman [1996]). In this sense, the predicted effect of demography should be viewed with care, most likely a lower bound of the true causal effect (as results in the next section suggest). Figure X shows the evolution of homicide victim rates by age group.

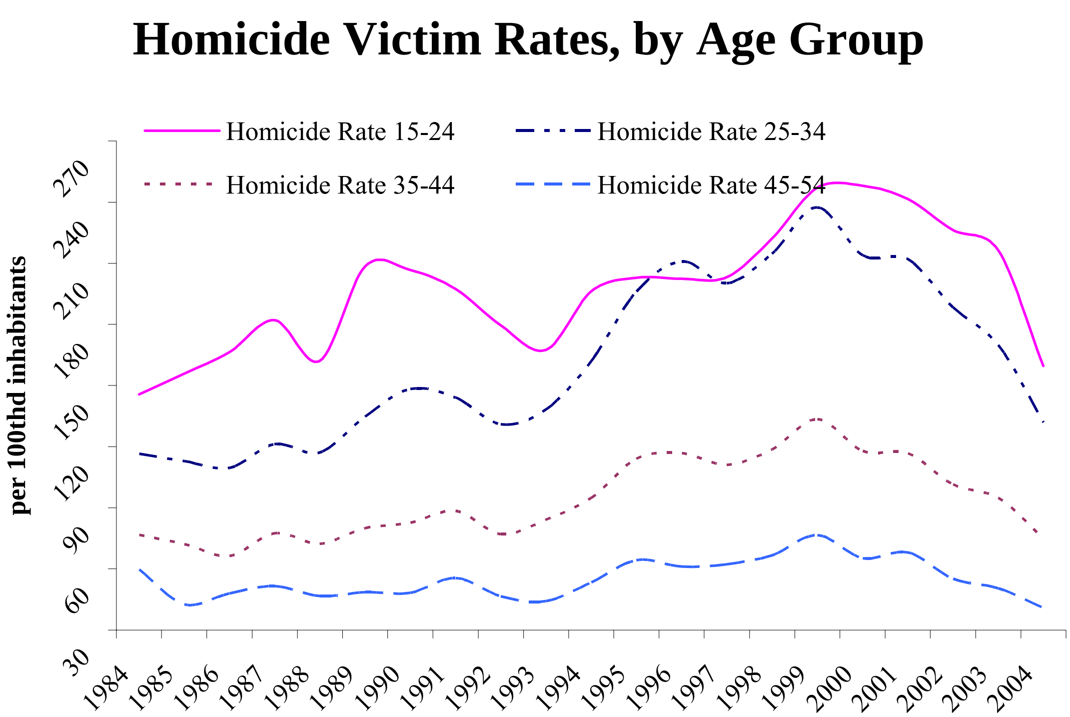


Figure IX Source: DATASUS

There seems to be a common component across age groups. However, movements in homicide rates are much more pronounced in the 15-24 and 25-34 groups. This suggests that most of the variation in homicide rates is explained by the variation in deaths among the more crime prone age groups, particularly the 15-24 group.

III. Panel Evidence

Ultimately demography and crime are not randomly determined but are choices of the agents. Consequently, the relationship between demography and homicides may suffer from the usual problems: of reverse causality and omission of common determinants. Additionally, using time series variation alone one cannot dismiss the possibility that the relation arising in figures VIII and IX, and table II are a product of a sheer coincidence (although demography and homicides seem related above and beyond the inclusion of a high-order polynomial of time).

Reverse causation seems highly improbable, at least as a first order phenomenon empirically. It is true that homicide victims are concentrated in the male age bracket 15-24. There are, however, too few murders to make a significant difference. For an illustration, at its 1999 peak, the city of São Paulo had 2418 homicides whose victims were 15-to-24-year-old males. Although the number is certainly very high, it amounts to no more than 0.25% of the 969,241 young males living in São Paulo that year. Furthermore, reverse causation, in this case, would bias the estimated relationship between demographics and crime towards zero

Demography has two pillars. One is fertility and mortality, which largely produced by decision made several years - if no decades - before. Second is migration, a shorter-term decision. Similarly, crime is a decision made in the present. From the first channel (fertility), there is little chance that demography and crime have a common cause. Migration is more challenging for the estimation strategy, and it is further discussed below.

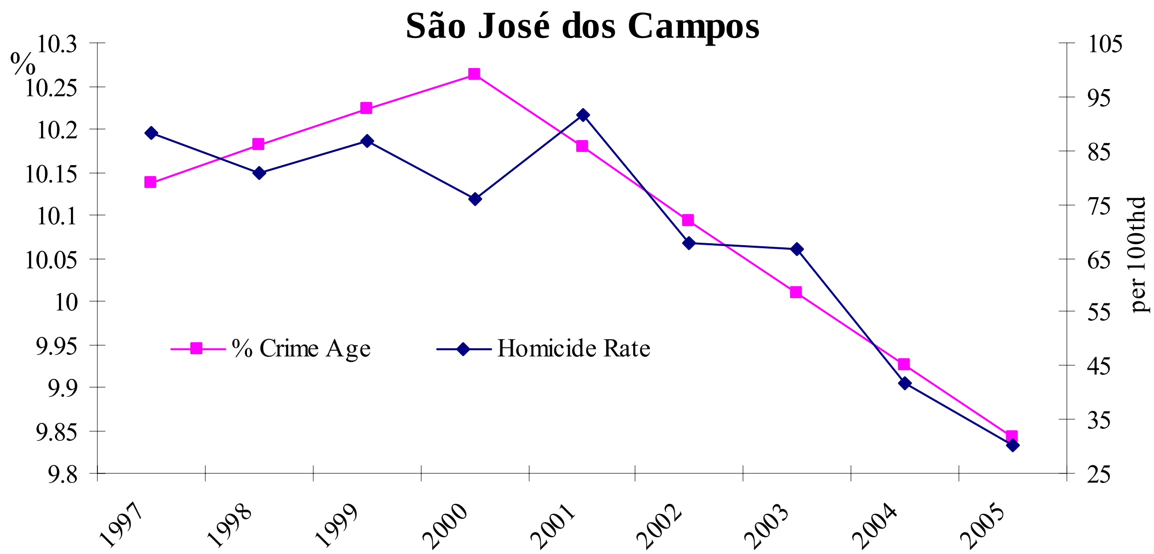
In this subsection we estimate how demography impacts violent crime using, along with variation over time, how demography evolved in different large cities in the state of São Paulo. Cross-city variation has the obvious advantage of allowing to evaluate the "coincidence" explanation. Let *i* be a city and *t* be a year. The estimated model is:

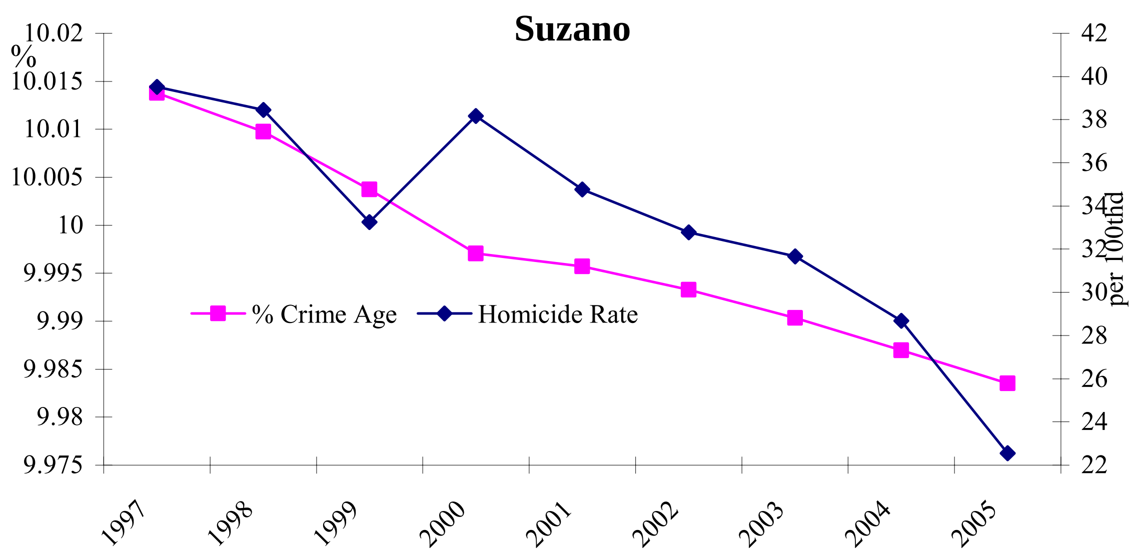
$$\begin{aligned} \log(Homicide)_{it} &= \beta_0 + \beta_1 \log(Male1524)_{it} + Controls_{it} + \\ &\sum_{t=1}^{T} \tau_t TIME_t + \sum_{i=1}^{I} \iota_i CITY_i + \varepsilon_{it} \end{aligned} \tag{1}$$

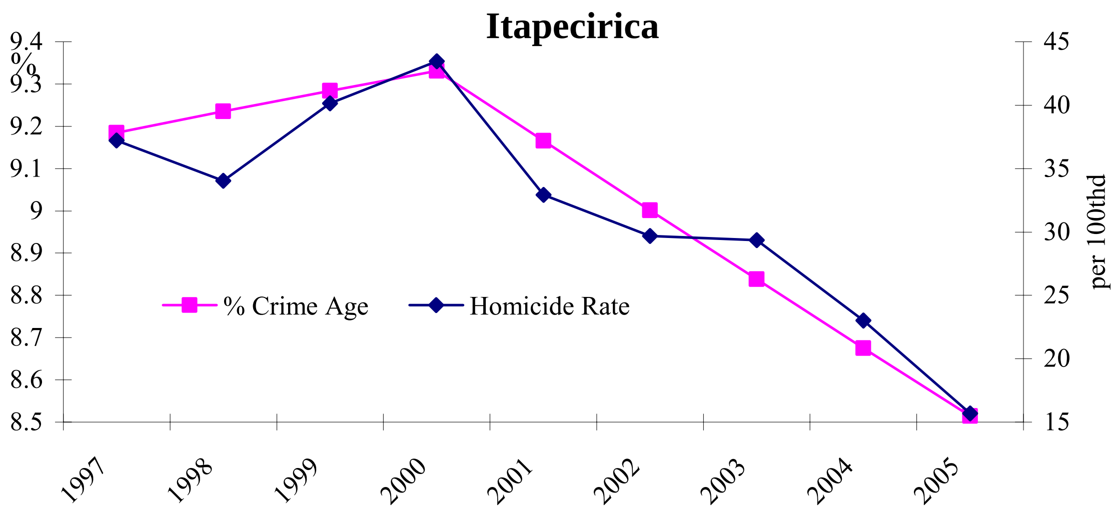
Homicides are rates per 100thd inhabitants, *Male*1524 is the percentage of 15-24 year-old males. $TIME_t$ is a full set of year dummies, and $CITY_i$ is a full set of city dummies. *Controls*, in some specifications, will include the log of population and the log of the high-school dropout rate. These two controls are quite important for our purposes. First, population will capture migration movements, the component of demography that is a product of current choices of agents. Second, it may be easier to maintain youngsters at school if there fewer of them.

With a panel structure, one can discard *all* pure time-series variation (and all pure cross-city variation), leaving only how demography changed *differently in different cities* as a source of variation to estimate its impact on homicides. Several more layers of coincidence are now necessary to produce the results spuriously. Second, we can account for all time-invariant heterogeneity among cities, which helps identifying the effect of demography. Figures IX-XIV illustrates graphically the type of variation explored when estimating equation (1). The proportion of 15-24 year-old males and homicide rates are depicted for five large cities statewide.[18]

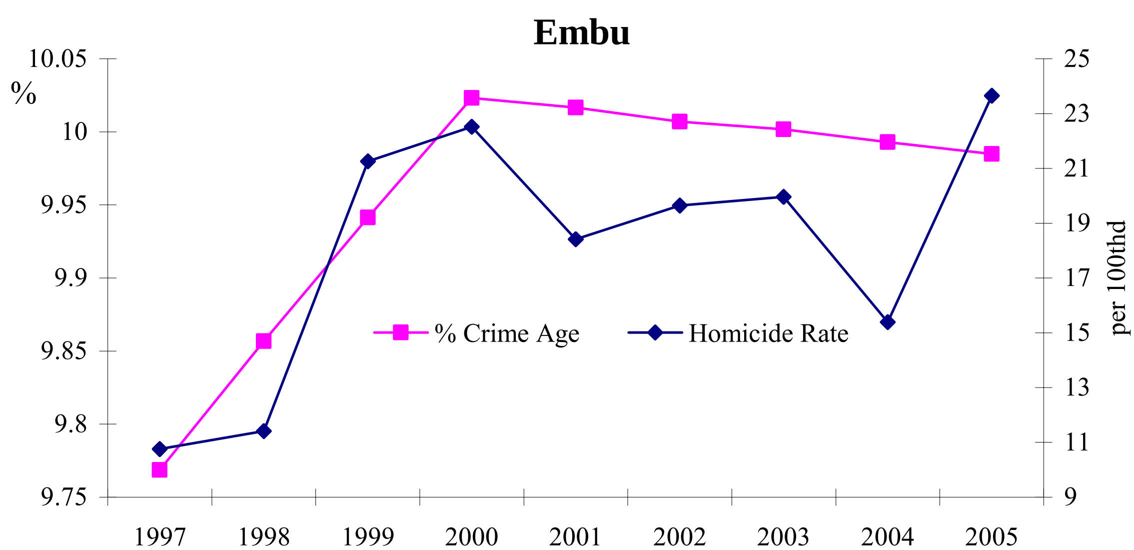
[18] We depict no more than five cities for the sake of conciseness. Several other cities would confirm the pattern.


São José dos Campos
%
10.3
10.25
10.2
10.15
10.1
10.05
10
9.95
9.9
9.85
9.8
105
95
85
75
65
55
45
35
25
per 100thd
% Crime Age
Homicide Rate
1997
1998
1999
2000
2001
2002
2003
2004
2005

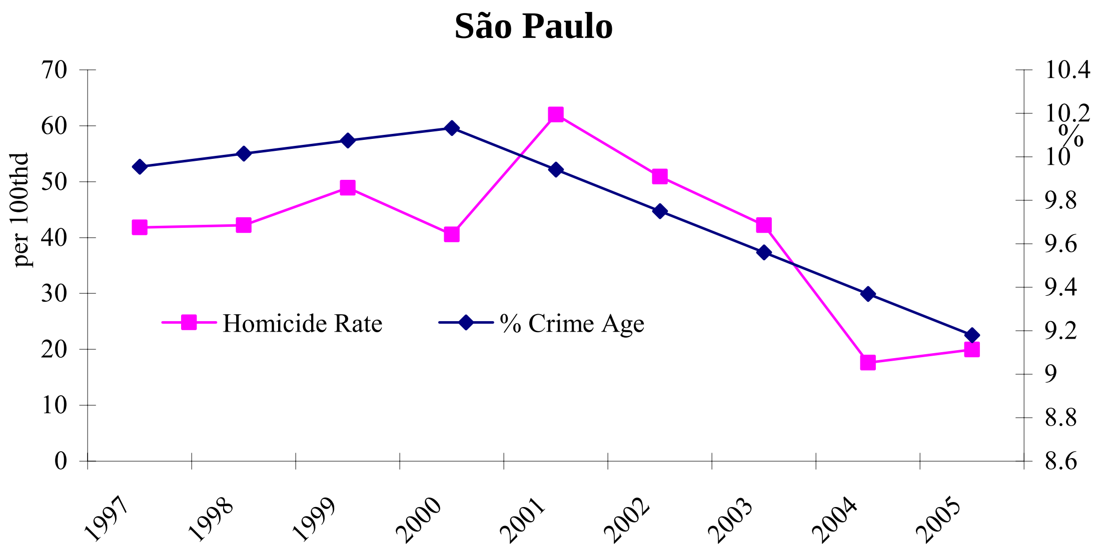

Suzano
%
10.02
10.015
10.01
10.005
10
9.995
9.99
9.985
9.98
9.975
42
40
38
36
34
32
30
28
26
24
22
per 100thd
% Crime Age
Homicide Rate
1997
1998
1999
2000
2001
2002
2003
2004
2005



Itapecirica
%
9.4
9.3
9.2
9.1
9
8.9
8.8
8.7
8.6
8.5
45
40
35
30
25
20
15
per 100thd
% Crime Age
Homicide Rate
1997
1998
1999
2000
2001
2002
2003
2004
2005



Embu
%
10.05
10
9.95
9.9
9.85
9.8
9.75
25
23
21
19
17
15
13
11
9
per 100thd
% Crime Age
Homicide Rate
1997
1998
1999
2000
2001
2002
2003
2004
2005



São Paulo, São José dos Campos and Itapecirica follow the typical pattern: proportion of 15-24 year-old males increased until 2000 and fell thereafter, with homicides following the same pattern. However, where the proportion of 15-24 year-old males increased less pronouncedly, and fell less pronouncedly (São Paulo), homicides also went up smoothly and fell relatively smoothly. In other cities, such as São José and Itapecirica, the proportion of 15-24 year-old males increased smoothly, but then drop steeply, homicides also increase smoothly and then dropped pronouncedly. Finally, consider Suzano and Embu. In the former, the proportion of 15-24 year-old males fell throughout the 1997-2005 period, and homicides followed the same pattern. On the other, in Embu the proportion of 15-24 year-old males increased sharply until 2000, and so did homicides; from 2000 onwards, the proportion of 15-24 year-old males fell very subtly, and so did homicides.

Table III present several models estimated using only data from the São Paulo Metropolitan Area (SPMA), in which case there is data starting in 1991.

TABLE III: Homicide Regression, SMPA 1991-2005
Dependent Variable: Log of Homicide Rate per 100thd inhabitants

	(1)[A]	(2)[A]	(3)[B]	(4)[B]	(5)[B]	(6)[B]	(7)[B]	(8)[B]	(9)[B]
Log(Crime Age)	3.92		4.95		2.02		1.71		
	(0.43)***		(0.59)***		(0.63)***		(0.70)**		
Log(Crime Age)$_{t-1}$		4.18		5.48		2.65		2.37	
		(0.46)***		(0.75)***		(0.77)***		(0.83)***	
Average Log(Crime Age)[A]									2.27
									(0.82)***
Log(Population)							0.19	0.18	0.14
							(0.28)	(0.31)	(0.32)
Log(Dropout Rate)[C]							0.04	0.03	0.04
							(0.06)	(0.06)	(0.06)
Year Dummies?	No	No	No	No	Yes	Yes	Yes	Yes	Yes
# Obs	515	482			515	482	514	481	481
R^2	0.12	0.13	0.70	0.70	0.76	0.77	0.76	0.77	0.77

All standard errors are White-Huber heteroskedastic corrected, unless otherwise noted
***: significant at the 1% level
**: significant at the 5% level
*: significant at the 10% level
A: OLS Regression
B: Fixed-Effects Regression
C: High-school dropout rate, moving average over the second and third lags
Source: Secretaria de Segurança Pública de São Paulo (SSP-SP), Secretaria de Estado da Educação (SEE) and Instituto Brasiliero de Geografia e Economia (IBGE)

Column (1) presents the simplest possible model, without city or year dummies and with no controls included. A 1% increase in the percentage of 15-24 year-old males is associated with 3.92% increase in homicides. To have a sense o practical importance, over the 1996-2000, there were 9.58% of young males in the SPMA, and 9.18% in the 2001-2005 period (see table I), which represents a 4.2% difference. Therefore, the coefficient implies that 16.37% of the reduction in homicides, which actually fell some 16% over the period.[19] Since it is not clear whether current or (recent) past demography matters, homicides are regressed on the lag of *Male*1524 (column (2)). Results are even stronger.

Interestingly, when city-fixed effects are accounted for, the impact of demography is *stronger*, for both contemporaneous and the lag (4.95% and 5.48%). As expected, including year dummies dampens results, and so does including the two controls. However, we can always reject the null hypothesis of that demography does not cause homicides at reasonable significant levels. The lowest possible estimates arise when the two controls, population and high-school drop-out rates. Neither population nor drop-out rates seem to belong to the equation, but the percentage of 15-24 year-old males do. At

[19] This is an extrapolation of a local interpretation of the log-in-log regression.

the lowest estimate, changes in the *Male*1524 imply a reduction of 7.14% in homicides from the second half of the 1990s and the first half of the 2000s. Finally, in column (9), the most complete model is estimated using a moving average of *MALE*1524. Results, as expected, are between those in columns (7) and (8). In table IV, again using data from the SPMA, several econometric robustness checks are performed. For the sake of conciseness all models are estimated with the moving average of column (9).

TABLE IV: Homicide Regression, SMPA 1991-2005

Dependent Variable: Log of Homicide Rate per 100thd inhabitants

	$(1)^{C}$	$(2)^{C}$	$(2)^{D}$	$(3)^{D}$	$(5)^{E}$	$(6)^{E}$	$(7)^{F}$	$(8)^{F}$
Average Log(Crime Age)^{A}	1.87	1.77	2.58	2.40	2.80	2.81	3.28	5.13
	(0.57)***	(0.67)***	(0.97)***	(1.10)**	(0.98)***	(1.07)***	(0.91)***	(1.07)***
Log(population)		0.01		0.01		0.04		-1.62
		(0.21)		(0.40)		(0.47)		(0.59)***
Log(Dropout Rate)^{B}		0.07		0.40		0.07		0.40
		(0.05)		(0.45)		(0.05)		(0.49)
ΔLog(Homicide Rate)$_{t-1}$					0.28	0.28		
					(0.05)***	(0.05)***		
Year Dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	481	481	482	481	436	436	317	316
R^2	0.79	0.79	0.76	0.76			0.83	0.83

All regressions account for city fixed effects. All standard errors are White-Huber heteroskedastic corrected, unless otherwise noted

***: significant at the 1% level

**: significant at the 5% level

*: significant at the 10% level

A: Moving average over current and first lag

B: High-school dropout rate, moving average over the second and third lags

C: Weighted Least Squares, Weight = Population½

D: GLS Regression, model for the error term: AR(1)

E: Arellano-Bond GMM Regression, all varibles in first-differences

F: Year ≥ 1997

Source: Secretaria de Segurança Pública de São Paulo (SSP-SP), Secretaria de Estado da Educação (SEE) and Instituto Brasiliero de Geografia e Economia (IBGE)

Four pairs of regressions are presented, including and not including the controls. All regressions account for city and year fixed effects. The first is a pair of Weighted Least Square regressions, in which the variance is modeled as a decreasing function of population. There are cities of wildly different sizes in the SPMA, and homicides are a relatively rare occurrence, observations from small cities are very noisy. Results are in line with those in column (9) table III. While in columns (1) and (2) the variance is modeled, in columns (3) and (4) the autocorrelation in modeled as an AR(1) process. Results are stronger.

In dynamic models, if the first lag belongs to the equation, and errors are auto-correlated (which columns (3) and (4) suggest is the case), the first lag is endogenous, and all coefficients are biased. To account for this possibility, we estimate the model by GMM using lags of the regressors as instruments (see Arellano and Bond [1991] for details). Even after including the first lag of homicide, results are once more stronger (columns (7) and (8)).[20]

Finally, the model is estimated with the sample restricted to 1997 onwards, for two reasons. First, by shortening the sample on the time dimension, one reduces the odds that time-varying unobserved heterogeneity across cities is driving results. Second, results on the 1997-2005 period are comparable with the other sample used, of large cities statewide. Results are significantly stronger than those with the whole 1991-2005 sample. Interestingly, increases in population seem to *reduce* crime.

In table V, some of the models presented in tables III and IV are replicated for a sample of all cities with more than 100,000 inhabitants (belonging or not to the SPMA).[21]

[20] The model is estimated in first-difference. After estimating the model by GMM, we tested for first and second order autocorrelation on the error term. This is important because if *second* order autocorrelation is still present, coefficients could be biased if the second lag of homicides belongs to the equation. While first order autocorrelation is indeed present, after including the first lag of homicides second order correlation is not present.

[21] We do not present all models for the sake of brevity. All other results are in line with those presented.

TABLE V: Homicide Regression, Large Cities 1997-2005
Dependent Variable: Log of Homicide Rate per 100thd inhabitants

	(1)^{A}	(2)^{B}	(3)^{B}	(4)^{B}	(5)^{B}	(6)^{BC}	(7)^{BD}	(8)^{BE}	(9)^{BE}	(10)^{BE}
Log(Crime Age)	6.06	6.32	4.67	6.37		5.91	6.03	5.80	3.12	
	(0.50)***	(0.58)***	(1.01)***	(1.07)***		(0.87)***	(1.24)***	(1.54)***	(1.83)*	
Average Log(Crime Age)					7.70					4.35
					(1.22)***					(1.77)**
Log(population)				-2.09	-2.97	-1.71	-1.88	-2.96	-2.99	-3.20
				(0.59)***	(0.70)***	(0.51)***	(0.68)***	(0.82)***	(1.10)***	(1.05)***
Log(Dropout Rate)				0.06	0.09	0.06	0.07	0.10	0.14	0.15
				(0.06)	(0.06)	(0.06)	(0.05)	(0.07)	(0.07)**	(0.07)**
ΔLog(Homicide Rate)t - 1								0.43	0.48	0.46
								(0.08)***	(0.09)***	(0.08)***
ΔLog(Homicide Rate)t - 2									0.37	0.36
									(0.07)***	(0.07)***
Year Dummies?	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	585	585	585	585	520	436	585	455	390	390
R^2	0.17	0.85	0.86	0.87	0.87	0.88	0.87			

All regressions account for city fixed effects. All standard errors are White-Huber heteroskedastic corrected, unless otherwise noted. Average Log(Crime Age) and Drop-outs as defined in tables III and IV.
***: significant at the 1% level
**: significant at the 5% level
*: significant at the 10% level
A: Simple OLS
B: City-fixed effects included
C: Weighted Least Squares, Weight = Population½
D: GLS Regression, model for the error term: AR(1)
E: Arellano-Bond GMM Regression, all varibles in first-differences
Source: Secretaria de Segurança Pública de São Paulo (SSP-SP) and Instituto Brasiliero de Geografia e Economia (IBGE)

Three important messages arise from the table V. First, when only the 1997-2005 period is considered results are stronger than with the whole 1991-2005 period. The inclusion of large cities outside the SPMA does not change results in any significant way. We interpret these results as suggesting two facts. First, the phenomenon is wider than the SPMA. Second, with a shorter time length it is less likely that demography captures time-varying heterogeneity. Finally, confirm previous estimates, population seem to reduce homicides, and some models now suggest that more high-school drop-out rates increase homicide as one should expect. The coefficient on population is puzzling in the light of the previous literature on social interaction that would predict crime rates are higher in larger cities (Glaeser, Sacerdote and Scheinkman [1996]). Our results suggest that, perhaps, economically dynamic and safer cities receive an influx of population. Since young males are more prone to moving, unobserved migration movements (imperfectly captured by population) would work towards biasing the impact of young males on homicides *towards zero*.

IV. Discussion: Alternative Explanations

There are other contributing factors to explain the radical shift in the trend of homicides in the state of São Paulo in general, and in the SPMA in particular. As one would expect, there was a policy reaction to the rise in homicides during the 1990s, at the municipal, state and national levels. In table VI, the most relevant policies adopted from the second half of the 1990s through 2007 are listed.

Table VI: Policy Interventions

Policy	Level	Date
Creation of INFOCRIM, a database system of crime georeferencig	SPMA and large cities	1999
Executive order likining number of police officer to INFOCRIM	SPMA and large cities	2007
Creation of DISQUE-DENÚNCIA, a anonymous crime hotline to deounce crime	State	2000
Creation of FOTOCRIM, a database of pictures of wanted and in prison	State	1999
Effective implementation of FOTOCRIM as an instrument of photograph identification	State	2002
Elaboration of the Plano de Combate aos Homicídios, with emphasis on capturing repeated murderers	State	2001
Adoption of "Dry Laws", legislation restrciting the recreational sales of alcohol	SPMA[A]	2001-2004[B]
Creation of a Municipal Police Forces	SPMA[C]	1962-2003[D]
Lei do Desarmamento ("Disarmament Law")	National	2003
"Operação Saturação", a centralized, systematic and permanent operation in drug-trafficking areas	State	2006

[A]: Adoption occurred in 16 out of the 39 cities that form the SPMA
[B]: Average adoption period is July-2002
[C]: 26 out of the 39 cities that form the SPMA have Municipal Police Forces
[D]: 7 Municipal Police Forces were created after 1998
Source: Kanh and Zanetic [2005] and Biderman, De Mello and Schneider [2007]

Most likely, all these policy interventions have merits of their own. The presence of a *Compustat* system such as INFOCRIM surely helps law enforcement. It is quite conceivable that stricter gun controls will prevent many silly homicides.

Nevertheless, it is difficult to imagine that policy interventions can account for the dynamics of homicide in the state of São Paulo. One reason is timing. All policy intervention occurred either exactly when the trend in homicides have been reversed (1999-2000), or afterwards. Although INFOCRIM was implemented in 1999, it became fully operational only in 2001, and data has not been used to determine local police force until now. The first implementation of dry laws was in 2001; municipal police forces have been created throughout the 1990s.

It is quite conceivable that these policies have contributed to the decline in homicides in the 2000s, although dry laws are the only policy for which there is hard evidence of effectiveness (Biderman, De Mello and Schneider [2007]). Their absence cannot account for the previous rise in 1990s, nor can they explain the reversal of the trend.

V. Conclusion

We find that the age structure explain a significant part of the variation in homicides at state of São Paulo over the last 16 years. The reversal of homicides rates were of the same magnitude of that in the 1980-1990 period in large US cities, most prominently New York.

These results are important *per se* for criminology science. While there is undisputable evidence that offenders are mostly for an specific age-gender group (males between 15 and 24 years old), whether age structure would appear at the aggregate level is not clear. Levitt [1999] finds a limited role for demography. Our result can reconcile these two seemingly opposing pieces of evidence. Although age structure must mechanically contribute to crime rates, whether it makes an aggregate difference will probably depend on law enforcement, the efficacy of the judicial system, institutional development, educational and labor opportunity for young males, etc. Perhaps, in the state of São Paulo, differently from the US, the environment was ripe for demography to "flourish" as a cause of homicides.

As for policy, these results are also important. While it is hard to influence demography in the short-run, sub-optimal policy over-reaction can be avoided if demography is the driving factor.

V. References

ARELLANO, M. AND BOND, S. “Some Specification Tests for Panel Data: Monte Carlo Evidence and an Application to Employment Equations,” *The Review of Economic Studies*, 58(1991), 277-297.

BIDERMAN, C., DE MELLO, J. AND SCHNEIDER, A. “Dry Law and Homicides: Evidence from the São Paulo Metropolitan Area,” Departamento de Economia, PUC-Rio: Texto para Discussão No 518, 2006

BLUMSTEIN, A. “Prisons,” *in* WILSON AND PETERSILIA, eds., *Crime*, San Francisco: ICS Press.

COOK, P. AND LAUB, J. “The Unprecedented Epidemic in Youth Violence,” *in* MOORE AND TONRY, eds., *Crime and Justice*: a Review of Research, Chicago: the University of Chicago Press, 1998.

DONOHUE, J. AND LEVITT, S., “The Impact of Legalized Abortion on Crime,” *The Quarterly Journal of Economics* 116 (2001), 379-420.

GLAESER, E., SACERDOTE, B. AND SCHEINKMAN, J., “Crime and Social Interaction,” *The Quarterly Journal of Economics* 111 (1996), 507-48.

GORING, C. *The English Convict*, Montclair: Patterson Smith, 1913.

LEVITT, S., “The Limited Role of Changing Age Structure in Explaining Aggregate Crime Rates,” *Criminology* 37 (1999), 581-97.

WILSON, J. AND HERRNSTEIN, R. *Crime and Human Nature*, New York: Simon & Schuster, 1985.