

Desesperança de vida: homicídio em Minas Gerais, Rio de Janeiro e São Paulo no período 1981/97*

Mônica Viegas Andrade**
Marcos de Barros Lisboa***

1 - Introdução

Nos últimos 20 anos a violência cresceu assustadoramente nos Estados do Rio de Janeiro e São Paulo, tornando-se, inclusive, a principal causa de mortalidade para os homens entre 15 e 44 anos de idade. Em São Paulo, o número de homicídios por 100 mil habitantes passou de 54,4 em 1981 para 128,4 em 1995, no grupo dos homens jovens entre 15 e 24 anos, e de 49,3 para 106,2 para os homens entre 25 e 44 anos de idade.¹ Esse aumento representa um crescimento de 136% e 115%, respectivamente. O Rio de Janeiro apresenta taxas de homicídio por 100 mil habitantes ainda maiores nesse período, embora o crescimento da violência não tenha sido tão elevado quanto o observado em São Paulo. Para os homens entre 15 e 24 anos, a taxa específica de mortalidade passa de 148,9 para 275,3 entre 1981 e 1995, significando um crescimento de 85%.

Em Andrade e Lisboa (2000) é calculado o número de anos de vida perdidos para diversas causas de mortalidade, que estima quantos anos, em média, um indivíduo viveria a mais caso cada causa de mortalidade não existisse. O resultado surpreendente é o crescimento relativo da violência como causa de mortalidade, contrastando com as demais causas que apresentam, na sua grande maioria, uma tendência de convergência aos índices dos países desenvolvidos, embora ainda estejam em níveis mais elevados. No começo da década de 90 a violência aparece como a principal responsável pela perda de anos de vida para os homens no Estado do Rio de Janeiro, e a segunda em São Paulo, sendo superada,

* Naércio Aquino e Carlos Martins comentaram com a generosidade habitual uma versão preliminar deste artigo. Os erros remanescentes são de nossa responsabilidade.

** Da EPGE/FGV-Rio e do Cedeplar/UFMG.

*** Da EPGE/FGV.

1 Dados apresentados em Andrade e Lisboa (2000).

apenas, pela mortalidade infantil. Entre 1981 e 1995, o número médio de anos de vida perdidos por cada homem residente no Estado do Rio de Janeiro passou de 1,57 para 3,42. Isso significa que se o problema do homicídio fosse erradicado cada homem viveria em média 3,42 anos a mais. A importância dos homicídios na mortalidade entre os homens nos levou a investigar, em maior detalhe, sua evolução ao longo do tempo e sua possível relação com fatores econômicos.

O objetivo deste trabalho é analisar a evolução da taxa de homicídio nos Estados de Minas Gerais, Rio de Janeiro e São Paulo entre 1981 e 1997. Em particular, procuramos verificar a existência de relação entre essa evolução e variáveis econômicas como salário real, desemprego, coeficiente de Gini, entre outras. Em que medida a tendência de crescimento da taxa de homicídio está associada à crise econômica das décadas de 80 e 90? É possível identificar algum padrão recorrente nos dados de homicídio?

A relação entre variáveis econômicas e violência tem sido objeto de diversos trabalhos na literatura econômica aplicada que, muitas vezes, utiliza como medida da violência a taxa de homicídios por 100 mil habitantes. Essa mensuração, entretanto, não nos parece adequada por três motivos principais. Em primeiro lugar, a taxa de homicídio varia consideravelmente entre as idades e os sexos. Em geral, as principais vítimas de homicídios são os homens jovens, com idades entre 15 e 30 anos. Dessa forma, as taxas de homicídios por 100 mil habitantes podem variar significativamente entre regiões em decorrência apenas da diferença nas composições etária ou de sexo. Por exemplo, a frequência de mortes entre os homens com idade entre 15 e 24 anos foi cerca de três vezes maior em 1981 no Rio de Janeiro do que em São Paulo (149,8 e 54,4, respectivamente), enquanto a taxa de homicídio por 100 mil habitantes nesse mesmo período foi apenas duas vezes superior (30,64 e 15,31). Esse resultado reflete a maior idade média e maior participação das mulheres na população do Estado do Rio de Janeiro em relação ao Estado de São Paulo. Essa diferença de magnitude pode resultar em viés nas análises controladas da evolução da taxa de homicídio, tratando-se de um caso típico de omissão de variáveis.

Em segundo lugar, a relação entre homicídio e variáveis econômicas pode mudar ao longo do ciclo de vida. Diversas atividades legais apresentam retornos crescentes com a especialização decorrente do seu exercício repetido. Mesmo que os retornos não sejam crescentes, a participação prévia do trabalhador no mercado de trabalho pode ser utilizada pelas empresas como um sinal sobre a sua qualidade ou produtividade. Além disso, o exercício de atividades ilegais no passado pode comprometer o acesso ao mercado de trabalho legal. Dessa forma, indivíduos com experiência no mercado legal de trabalho podem ter acesso a remunerações e empregos maiores do que os que exercem atividades ilegais, e essa diferença pode aumentar com o ciclo de vida.

É possível, portanto, que o impacto das variáveis econômicas sobre a parcela da população que se dedica a atividades ilegais seja diferenciado ao longo do ciclo de vida. Para um jovem, a diferença entre as atividades legais e ilegais pode

ser menor do que para um homem mais velho que participa do mercado de trabalho legal. Nesse caso, as variáveis econômicas podem ter um impacto maior sobre os jovens do que sobre os mais velhos. Caso a frequência de homicídios aumente com a parcela da população envolvida com atividades ilegais, o impacto das variáveis econômicas sobre a taxa de homicídio será, igualmente, diferenciado ao longo do ciclo de vida.

Em terceiro lugar, o acesso diferenciado ao mercado de trabalho entre a população que participa do mercado legal de trabalho e a que participa de atividades ilegais pode resultar em efeitos geracionais sobre a taxa de homicídios (*efeito inércia*). Suponhamos, uma vez mais, que os homicídios cresçam com a parcela da população dedicada às atividades ilegais. Caso o retorno às atividades legais seja custoso, uma geração que quando jovem tem maior parcela dedicada às atividades ilegais tende a manter essa maior parcela ao longo de todo o ciclo de vida e, portanto, uma taxa maior de homicídio. Isso significa a possibilidade de existência de um efeito inércia dos homicídios em cada geração. Em particular, a taxa de homicídios em um ano pode ser elevada não porque as variáveis econômicas apresentam certo comportamento naquele ano, mas porque esse comportamento foi observado quando a geração era jovem, sendo seus efeitos diluídos ao longo do ciclo de vida da geração.

A eventual existência do efeito inércia pode ser testada desagregando-se os dados de homicídio por cada grupo etário distinto e acompanhando esses dados para cada geração por vários anos. A existência de autocorrelação na série de taxas de homicídio por geração, quando controlada pelas demais variáveis, é um indicador da possível existência do efeito inércia. Nesse caso, a taxa de homicídio de uma dada geração em um certo ano seria uma das variáveis relevantes para prever a taxa da mesma geração no ano seguinte. Uma geração violenta quando jovem tenderia a apresentar maiores índices de violência por todo o seu ciclo de vida.

Neste trabalho, a violência é mensurada através da construção de frequências de morte por homicídio para cada idade, sexo, ano e região de residência. A frequência é construída a partir dos dados de homicídios divididos pela população residente em cada região condicionados na idade, sexo e ano. A construção dessa base de dados nos permite estimar a relação entre probabilidade de morte por homicídio e ciclos econômicos para cada idade específica. Como o fenômeno da violência é concentrado na população masculina em idade ativa, calculamos a probabilidade de morte para os homens dos 15 aos 40 anos de idade. Além disso, esse tratamento dos dados nos permitiu construir a base de dados segundo coortes,² em que cada uma foi definida pelo ano em que os homens têm 15 anos de idade.³

2 O conceito de coorte segue a definição proposta por Ryder (1965), ou seja, um conjunto de indivíduos que experimentou os mesmos eventos dentro de um mesmo intervalo de tempo.

3 Poucos trabalhos na literatura analisam os dados em coorte em função da dificuldade de disponibilidade de pesquisas desse tipo. Os trabalhos de Tauchen e Witt (1994), e Stenfensmeier, Streifel, Shihadeh (1992) analisam dados em coorte.

Existem basicamente três tipos de abordagem empírica que tentam explicar a relação de causalidade entre violência e condições socioeconômicas: análises em *cross-sections*, em séries de tempo e trabalhos baseados em pesquisas de “vitimização” que acompanham o comportamento individual. Os resultados observados nesses três tipos de análises são bastante diferenciados, sobretudo no que se refere ao impacto do desemprego. Esses resultados são discutidos na quinta seção e comparados com os obtidos na nossa análise.⁴

Neste trabalho utilizamos um *mix* de *cross-section* com série de tempo já que acompanhamos os resultados de morte por homicídio para três estados brasileiros — Rio de Janeiro, São Paulo e Minas Gerais — durante o período 1981/97 com dados anuais. A análise em séries de tempo é privilegiada por ser mais adequada para se estudar a relação entre ciclos econômicos e violência e também o efeito inércia. De fato, uma das dificuldades com análises em *cross-section* é a possível ocorrência de correlações espúrias. A possibilidade de migração entre diversos estados sugere que o bem-estar para indivíduos semelhantes em regiões distintas deve ser igualmente semelhante, caso contrário estes poderiam simplesmente emigrar para a região com maior bem-estar. Por essa razão, regiões com maiores taxas de homicídio devem oferecer aos seus habitantes outras compensações de bem-estar tais como maiores salários reais ou melhor acesso a outros bens públicos que não segurança. Nesse caso, a existência de correlação positiva em uma base de dados *cross-section* — por exemplo, entre salário real e violência — pode não significar que aumentos do salário real aumentem a violência, mas apenas que o custo de bem-estar decorrente da maior violência em uma região é compensado pelo maior salário real, evitando a ocorrência de migração. Inclusive, é possível que um aumento do salário real simultâneo nas diversas regiões tenha como consequência a redução da violência. A construção de uma base de dados em painel pode permitir observar flutuações concomitantes das possíveis variáveis explicativas da taxa de homicídio em todas as regiões, evitando a ocorrência de correlações espúrias. Esse ponto é retomado na Seção 5, quando comparamos nossos resultados com a literatura existente.

O método de estimação utilizado neste trabalho é uma generalização do método do Mínimo Qui-Quadrado aplicado ao modelo Logit de Berckson [ver Maddala (1983) e Amemya (1985)]. Esse método consiste da estimação de um modelo logístico para variáveis qualitativas quando os dados estão disponíveis na forma de frequências. No caso específico deste trabalho os dados estão agrupados segundo idade, região de residência e ano de ocorrência do homicídio. A variável dependente é a probabilidade de morte por homicídio em cada idade, região e ano, obtida pelos dados do sistema de informações sobre mortalidade disponibilizados pelo Datasus. Como variáveis independentes, utilizamos indicadores

4 Freeman (1994) faz uma síntese dos resultados encontrados para a economia americana nos três tipos de abordagem.

socioeconômicos de cada região e atributos específicos da geração. Os dados foram estimados em duas bases. A primeira, denominada base completa, contém 1.410 células. Na segunda estimação, construímos uma para cada idade, com modelos estimados para 25 sub-bases, uma para cada idade dos 15 aos 40 anos.

Para estudar a possibilidade do efeito inércia utilizamos como variável independente em cada ano a probabilidade de mortes por homicídio da mesma geração no ano anterior. Esse tratamento implica, porém, a alteração da fórmula da variância do tradicional modelo de Berckson e requer uma correção distinta da heterocedasticidade. O método e o modelo estimado estão apresentados em detalhes na Seção 3.

O trabalho apresenta basicamente três resultados principais. Em primeiro lugar, a organização dos dados segundo as coortes dos indivíduos parece ser um tratamento bastante adequado no entendimento dos ciclos de violência. A observação da evolução da probabilidade de morte segundo as coortes descrita nos Gráficos 1, 2, 3 e 4 dá robustez à hipótese de persistência do ciclo de violência. O padrão de comportamento entre as coortes é praticamente o mesmo: a probabilidade de morte é crescente com a idade, atingindo, na grande maioria, o pico na faixa dos 20 aos 25 anos, quando o movimento se inverte e se torna decrescente com a idade.⁵ O ciclo de cada coorte é de aproximadamente 25 anos. Esse resultado é pronunciado com maior nitidez no Rio de Janeiro, mas o padrão se repete para os outros dois estados. Para o Estado do Rio de Janeiro apresentamos os Gráficos 1 e 2. No primeiro, mostramos a evolução das coortes que tinham 15 anos em 1982, 1983 e 1984, respectivamente. Como a base de dados disponível só nos permite obter informações até o ano de 1997, o período máximo que podemos acompanhar essas coortes é de 16 anos. No segundo gráfico, apresentamos as coortes que tinham 15 anos no período 1972/76, evidenciando o ritmo de decrescimento da probabilidade de morte com a idade. A construção dos dados em termos de probabilidade para cada idade parece sugerir que a taxa de homicídio em cada é ano é a soma de ondas de violência que se superpõem, em que cada onda corresponde a uma geração específica.

O segundo resultado do trabalho se refere à relevância das variáveis econômicas para explicar a taxa de homicídio. As variáveis econômicas parecem ser relevantes para os jovens entre 15 e 19 anos. A partir dos 20 anos de idade, a variável mais importante para explicar a violência é o componente de inércia mensurado neste trabalho pela inclusão da probabilidade defasada. Esse resultado difere dos trabalhos já existentes na literatura na medida em que incorpora o efeito coorte. A probabilidade defasada é a de um indivíduo da mesma coorte ter sido assassinado no ano anterior.

5 A relação entre taxa de participação no crime e idade é levantada por vários autores e parece ser robusta para qualquer tipo de crime. Grogger (1997) justifica esse padrão por meio do comportamento dos salários que são crescentes com a idade à medida que os indivíduos adquirem maior experiência. Se o comportamento do crime é sensível aos salários, quando os salários se elevam a taxa de criminalidade cai, pois o custo de oportunidade da atividade criminosa se eleva.

Gráfico 1

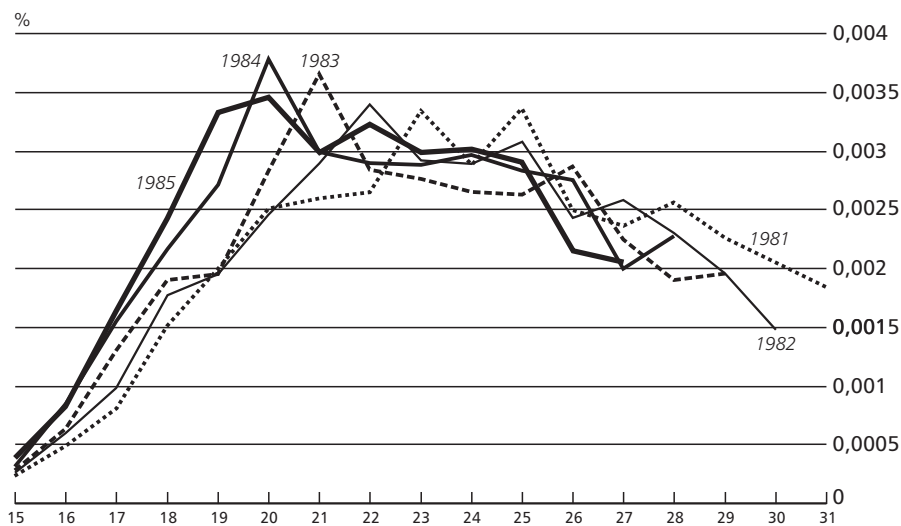
Evolução da taxa de homicídios por gerações - Rio de Janeiro

Gráfico 2

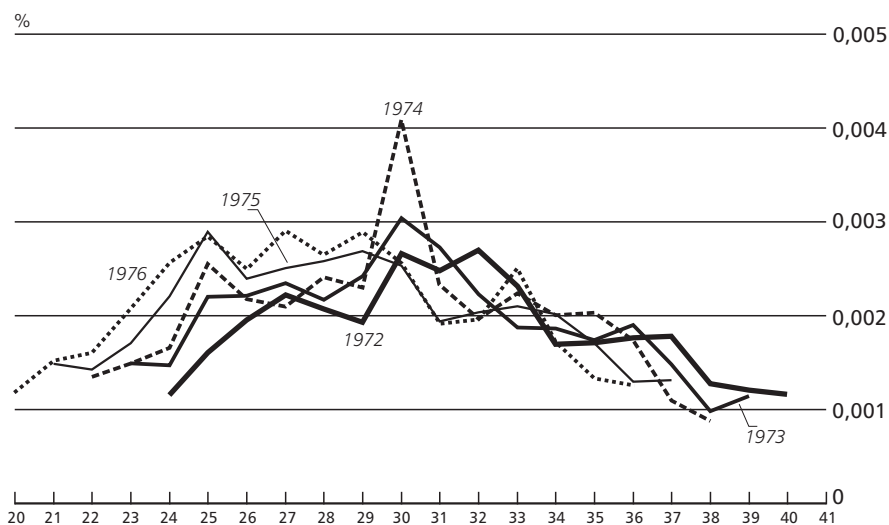
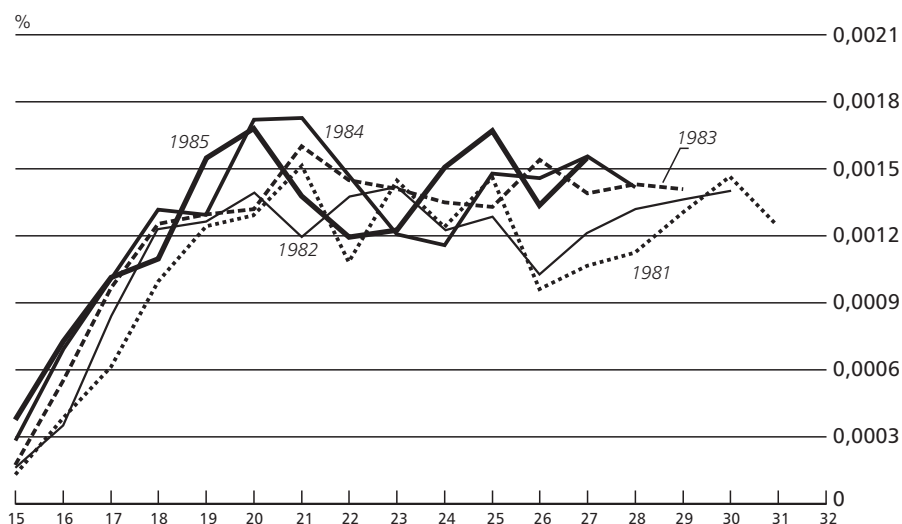
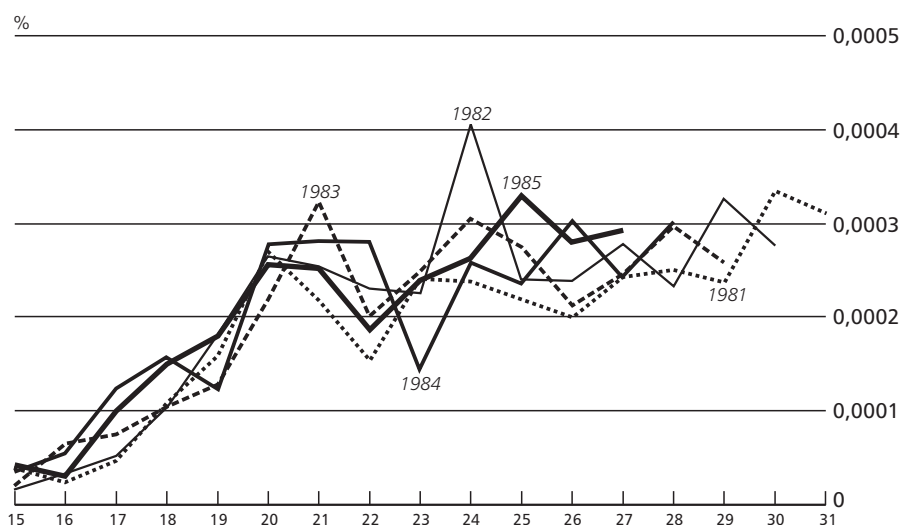
Evolução da taxa de homicídios por gerações - Rio de Janeiro

Gráfico 3

Evolução da taxa de homicídios por gerações - São Paulo

Desesperança de vida: homicídio em Minas Gerais, Rio de Janeiro e São Paulo no período 1981/97

Gráfico 4

Evolução da taxa de homicídios por gerações - Minas Gerais

Uma possível interpretação desse resultado relaciona, como já sugerimos, a idade da geração e a parcela dedicada às atividades ilegais. Indivíduos mais jovens se moveriam das atividades legais para as ilegais com maior facilidade do que os mais velhos, e esse movimento seria influenciado pelas variáveis econômicas. O aumento da população dedicada às atividades ilegais, por sua vez, resultaria em maiores taxas de homicídio. Dessa forma, uma geração que quando jovem apresente uma parcela elevada dedicada às atividades ilegais tende a manter essa parcela elevada durante todo o ciclo de vida e, portanto, uma taxa maior de homicídio. Essa hipótese, porém, ainda está por ser testada. Permanece como resultado da nossa análise, no entanto, que a política pública de segurança talvez deva se focar na população jovem masculina e que o controle da taxa de homicídio nesse grupo tem impactos sobre a taxa de homicídio de toda a geração. Entretanto, os efeitos dessa política só podem ser percebidos no longo prazo.

Em terceiro lugar, a diferença entre as taxas de homicídio no Rio de Janeiro e São Paulo é significativamente reduzida quando essas taxas são controladas pelo efeito inércia e pelas variáveis econômicas. Para a maioria das idades, a diferença não-controlada das taxas de homicídio é duas vezes superior no Rio de Janeiro do que em São Paulo, enquanto na maioria das regressões essa diferença cai para cerca de 20%.

Provavelmente, a principal fragilidade da nossa análise é a não-inclusão de uma variável como *proxy* para o sistema de punição e a política de segurança. Os únicos dados disponíveis para uma série contínua são as ocorrências policiais correlacionadas com a própria taxa de criminalidade.⁶ Infelizmente, não conseguimos base de dados confiável que descrevesse o número de prisões por ano, o gasto com segurança, ou qualquer outro indicador da atividade policial. A inexistência de base de dados minimamente confiáveis talvez seja o principal problema da análise empírica da violência no Brasil, além de refletir, do nosso ponto de vista, um aparente descaso do Estado com qualquer política de segurança de longo prazo. Surpreendeu-nos, em particular, a inexistência, inclusive, de qualquer série temporal para o período analisado informando o número de presos em cada um dos três estados.

Nossos resultados parecem sugerir, porém, que as eventuais diferenças nas políticas de segurança talvez não tenham um impacto significativo nas diferenças entre as taxas de homicídio observadas no Rio de Janeiro e em São Paulo quanto o sugerido pelos dados não-controlados. Esses resultados, no entanto, devem ser tratados com cautela. Pode ser que a política de segurança esteja correlacionada com as variáveis econômicas dos estados, onde os mais ricos teriam maiores recursos para a execução de uma melhor política de segurança. Além

6 Em versão preliminar, tentamos incluir variáveis *dummies* associadas às eleições como aproximações para o sistema de segurança de cada governo. Entretanto, o número de observações nas bases por idade não é suficiente para se estimar com significância os coeficientes.

disso, as variáveis econômicas utilizadas não são específicas a cada coorte, mas sim a todas aquelas que vivem na mesma região em cada ano. Dessa forma, nossos resultados devem ser tratados como uma primeira abordagem, ainda preliminar, que deverá ser objeto de estudos adicionais sobre o dramático aumento das mortes por homicídios durante as décadas de 80 e 90.

Este trabalho contém cinco seções adicionais. A próxima seção descreve a base de dados utilizada e as principais variáveis incluídas no modelo. A terceira seção discute a metodologia empregada. Na quarta seção apresentamos os principais resultados encontrados na estimação do modelo. Na quinta seção discutimos alguns resultados existentes na literatura e algumas estimações feitas para o caso brasileiro confrontando-os com os resultados empíricos encontrados neste trabalho. A sexta seção discute possíveis extensões do trabalho.

2 - A construção da base de dados

2.1 - Séries utilizadas

Neste trabalho, a variável utilizada para mensurar a taxa de criminalidade é a frequência de mortes por homicídio. Para cada ano e região dividimos o total de homicídios para cada idade pela população masculina residente com a mesma idade. Os dados de população foram obtidos a partir dos Censos de 1980 e 1991 e a Contagem de 1996. Para os demais anos, estimamos a população utilizando uma interpolação log-linear.

A fonte de dados de mortalidade utilizada é o Sistema de Informação de Mortalidade (SIM), disponibilizado pelo Ministério da Saúde por intermédio da Fundação Nacional de Saúde e do Datasus para o período 1979/97.⁷ A fonte de informação primária dessa base são os atestados de óbito emitidos pelos cartórios civis, que contêm informações sobre data do óbito, idade, sexo, estado civil, local de ocorrência, causa de mortalidade, área, bairro e município de residência, ocupação e instrução. Apesar da enorme gama de informações, essa base apresenta deficiência de algumas variáveis como educação, estado civil, ocupação, entre outras, que dificultam o seu uso. Neste trabalho utilizamos apenas as variáveis consideradas prioritárias pelo Ministério da Saúde — idade, sexo e causa de mortalidade — nas quais o índice de não-preenchimento chega em média a 7%. A causa de mortalidade está codificada até 1996 segundo a 9ª Classificação Internacional de Doenças, (CID9), e em 1997 de acordo com a 10ª revisão (CID10).

Nos Estados de São Paulo e Minas Gerais, o número de homicídios foi contabilizado utilizando-se as frequências observadas. No Estado do Rio de Janeiro, no entanto, para computar o número de homicídios, duas causas de mortalidade foram consideradas: o grupo “homicídios e lesões provocadas intencionalmente”,

7 Apesar de os dados de mortalidade estarem disponíveis a partir de 1979, os dois anos iniciais da pesquisa apresentam muitos problemas de tabulação dos dados.

e algumas das causas classificadas como “outras violências”. O Rio de Janeiro apresenta um elevado número de mortes por arma de fogo ou branca, porém classificadas como de intencionalidade ignorada. Enquanto essas mortes em São Paulo, por exemplo, jamais atingem 7% da mortalidade por homicídio dos homens entre 15 e 24 anos, no Rio de Janeiro em alguns anos esse número chega a 37%, apresentando, no entanto, um comportamento bastante errático.⁸ Esses resultados sugerem um problema com as bases de dados de atestado de óbito nesse estado.⁹ Para contornar essa dificuldade utilizamos o seguinte ajuste nos dados de homicídios: supusemos que a participação dos homicídios no total das causas *homicídios e outras violências* no Rio de Janeiro fosse, em cada ano, para cada idade e sexo, idêntica à obtida em São Paulo. A hipótese básica para esse procedimento é que as mortes acidentais correspondem a um evento aleatório de igual distribuição no Rio e em São Paulo, sendo porém correlacionadas com o nível de homicídios. Essa padronização, contendo a descrição das variáveis utilizadas, está detalhada no apêndice 1.

A escolha dos Estados de São Paulo, Minas Gerais e Rio de Janeiro se deve a dois motivos: em primeiro lugar, como já mencionado na Introdução, pela integração do mercado de trabalho desses três estados, e, em segundo, pelos menores índices de subenumeração de óbitos. Nas regiões Norte e Nordeste, por exemplo, os índices de subenumeração estimados pelo IBGE alcançam 50% em alguns estados.

A mensuração da taxa de criminalidade por meio da probabilidade de morte a cada idade permitiu construir a base de dados de duas formas. No primeiro caso construímos uma base para cada idade e em cada estado dos 15 aos 40 anos, que corresponde à faixa etária em que a violência é a principal causa de mortalidade. As bases de dados para cada idade específica contêm em média 42 células. Com isso, obtemos 25 bases de dados e podemos estimar coeficientes específicos para cada idade.

A segunda foi construída acompanhando os indivíduos de acordo com sua coorte, que é definida pelo ano em que os indivíduos tinham 15 anos. A primeira coorte corresponde aos indivíduos que tinham 15 anos em 1981. A primeira observação dessa coorte é a probabilidade de morte por homicídios dos homens de 15 anos em 1981, a segunda é a probabilidade de morte dos homens que tinham 16 anos em 1982, a terceira é a dos que tinham 17 anos em 1983, e assim sucessivamente até a probabilidade de morte dos homens aos 31 anos em 1997. A segunda coorte corresponde aos homens com 15 anos em 1982, e assim sucessivamente. Desse modo, foram construídas 43 coortes para cada estado, contabilizando 1.410 células. A organização das observações segundo coortes nos permitiu perceber o efeito inércia do ciclo de violência. Ademais, a probabilidade defasada foi construída incorporando o efeito coorte, já que se constitui da

8 Para a série temporal desses dados, ver Andrade e Lisboa (2000).

9 Esse problema é bem conhecido na literatura. Ver, por exemplo, Carneiro e Phebo (1999).

probabilidade de um indivíduo da mesma coorte ter morrido no ano anterior. Os Gráficos 1 a 4 ilustram com clareza a superposição dos ciclos de violência. Cada nova coorte tem um pequeno deslocamento, de modo que toda a geração se diferencia da anterior.

Vale ressaltar o *trade-off* existente na construção dessas duas bases de dados. Por um lado, a cada idade elas permitem inferir o impacto diferenciado das diversas variáveis independentes na probabilidade de homicídio. Em particular, podemos verificar em que medida esses coeficientes mudam significativamente com a idade. Essa abordagem, porém, reduz o número de células. A base de dados completa, embora nos permita obter resultados com maior grau de significância, requer a especificação de uma forma funcional prévia relacionando os diversos coeficientes das variáveis independentes com a idade da geração.¹⁰

As variáveis de controle utilizadas tentam, na medida do possível, incorporar as variáveis discutidas na literatura sobre economia do crime, que segue a contribuição de Becker (1968). Em decorrência do mau preenchimento dos dados, não podemos utilizar as variáveis dos próprios indivíduos contidas na base de dados de mortalidade, caso contrário incorreríamos em viés de seleção amostral. Diante disso, utilizamos as variáveis médias para cada região em cada ano.

As seguintes variáveis de controle foram incluídas: nível de escolaridade média da população economicamente ativa (PEA), coeficiente de Gini, taxa de desemprego, número de domicílios chefiados por mulheres, salário real médio da população ocupada, nível de preços, probabilidade defasada e duas variáveis *dummies*: uma para o Estado do Rio de Janeiro e outra para o Estado de São Paulo.¹¹ Essas variáveis são os controles usualmente utilizados nos trabalhos empíricos. O nível de escolaridade e o salário real seriam medidas dos retornos da atividade legal; a taxa de desemprego se associa a uma medida das oportunidades dos indivíduos no mercado de trabalho; a taxa de domicílios chefiados por mulheres é uma aproximação do grau de integração social e das mudanças sociais. O índice de desigualdade da renda descreve a posição relativa dos indivíduos. Um aumento da desigualdade, para os indivíduos mais pobres, faz com que a distância entre o retorno da atividade legal e da atividade do crime se eleve. Todas essas variáveis são condicionadas ao ano e à região.

A taxa de inflação foi incluída para captar possíveis distorções dos preços relativos decorrentes do processo inflacionário vivido nessas duas décadas que alterem os retornos das atividades, além das possíveis implicações decorrentes da maior volatilidade da renda real em períodos de inflação elevada. Ao contrário das demais variáveis, a taxa de inflação é condicionada apenas ao ano, sendo comum às três regiões.

10 Uma extensão possível desse trabalho seria utilizar uma abordagem não paramétrica para estimar essa forma funcional.

11 Os dados de probabilidade defasada foram construídos também por coortes. Witt e Witt (1998) mostram em um modelo de série de tempo que o aumento da participação feminina no mercado de trabalho está positivamente correlacionado com taxas de criminalidade mais elevadas.

As *dummies* para os Estados do Rio de Janeiro e São Paulo foram incluídas para verificar a existência de especificidade regional não-tratada pelas demais variáveis de controle.

A probabilidade defasada busca captar o efeito inércia¹² e é condicionada à geração e à região. Em cada ano, utilizamos como uma das variáveis de controle a frequência de homicídios da mesma geração no ano anterior. Considere, por exemplo, a geração que tinha 15 anos em 1985 no Rio de Janeiro. Assim, em 1990, quando esses indivíduos terão 20 anos, a probabilidade defasada corresponde à probabilidade de morte por homicídio dos homens que tinham 19 anos no Rio de Janeiro em 1989.

Devemos ressaltar que, lamentavelmente, como já discutimos na introdução, não utilizamos qualquer indicador das políticas de segurança que possam ter impacto na criminalidade e que talvez sejam distintas nos diversos estados e ao longo do período. Não conseguimos encontrar qualquer base de dados confiável contendo o número de prisões efetuadas em cada ano e região ou, mesmo, o total de presos existentes. O único dado disponível que encontramos foi o total de ocorrências, correlacionado com a própria atividade criminal e, portanto, não serve como indicador da política de segurança.

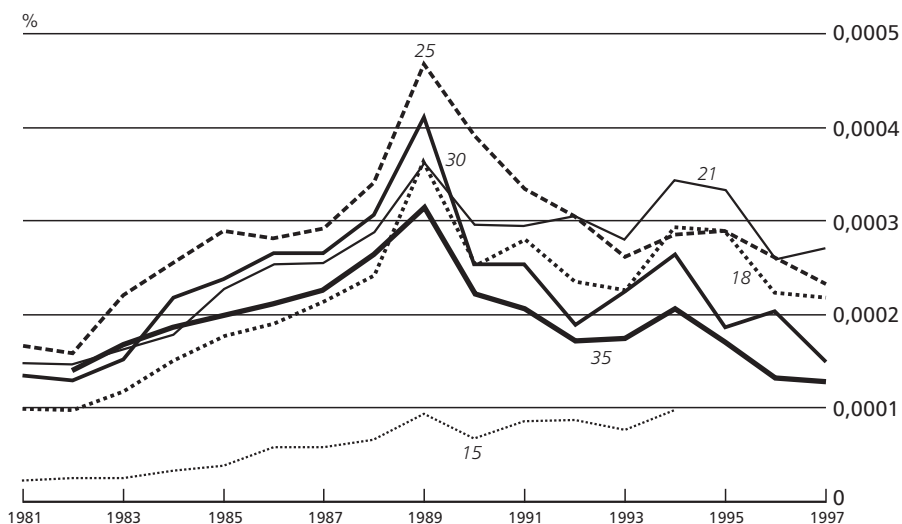
2.2 - Comportamento das variáveis

As frequências de homicídio não-controladas apresentam um comportamento distinto nos Estados do Rio de Janeiro e São Paulo em comparação com o Estado de Minas Gerais (Gráficos 5, 6 e 7). Nos dois primeiros, as frequências apresentam um comportamento relativamente similar nas diversas idades. No Estado do Rio de Janeiro esse comportamento é bastante nítido. Todas as idades, exceto 35 anos, apresentam um movimento ascendente da probabilidade de morte até o ano de 1989, quando ocorre o pico, seguido de uma longa inflexão dos anos 90 até 1993, pequena elevação em 1994 e tendência de queda a partir de 1995. No caso da probabilidade aos 35 anos, o movimento ascendente se inicia em 1989.

Em São Paulo o movimento é menos homogêneo, apresentando um padrão distinto para os grupos mais jovens. As probabilidades de 15, 18 e 21 anos apresentam movimento similar: elevação até o ano de 1987, queda em 1988, novo ciclo ascendente em 1990 e 1991, redução nos anos de 1992 e 1993 e tendência de elevação a partir de 1994. As idades de 25 e 30 anos apresentam elevação até o ano de 1983 seguida de estabilidade das taxas até 1986, pequena elevação em 1987, redução novamente em 1988, novo ciclo ascendente em 1989 e 1990, retração de 1991 a 1993 e tendência de elevação a partir de 1994. A comparação entre os dois Estados sugere um comportamento mais cíclico no Estado de São Paulo e tendência de elevação da violência no final dos anos 90, contrária

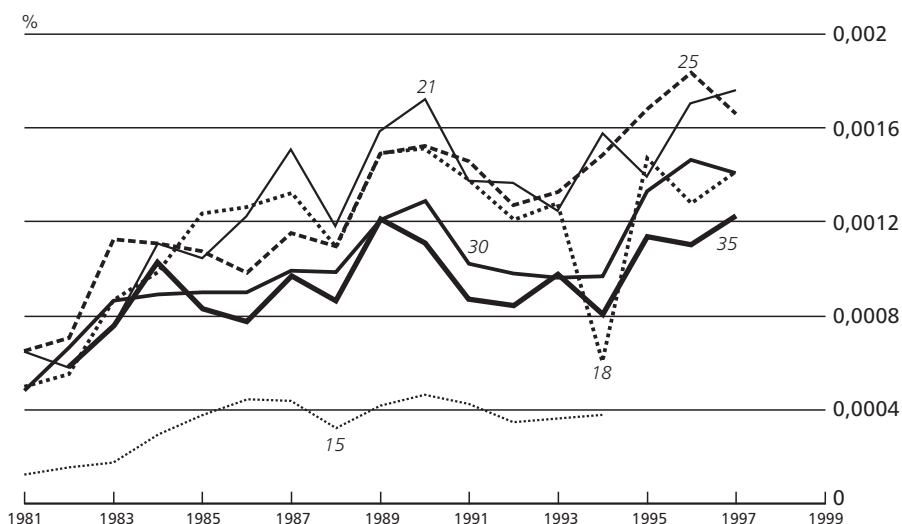
12 Fajnzylber, Lenderman e Loayza (1999) também utilizam como controle a taxa de criminalidade defasada.

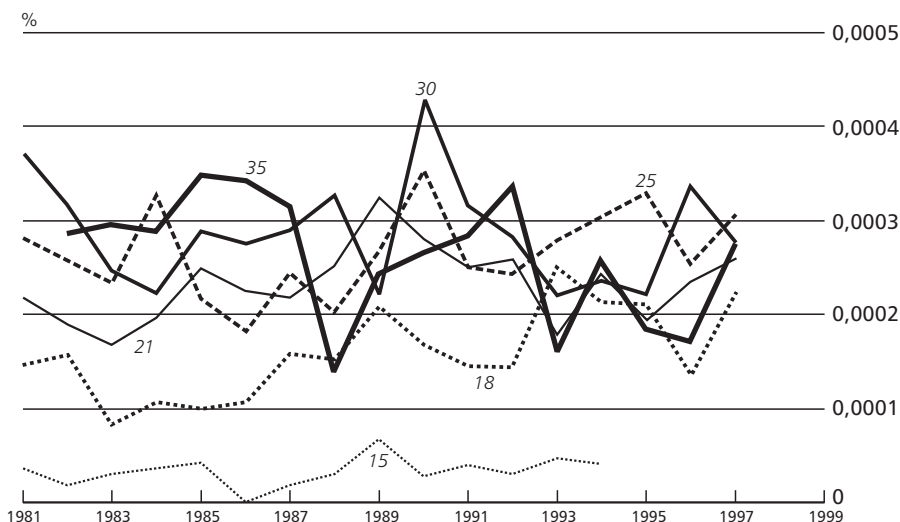
Gráfico 5

Evolução da taxa de homicídios por idade - Rio de Janeiro

Desesperança de vida: homicídio em Minas Gerais, Rio de Janeiro e São Paulo no período 1981/97

Gráfico 6

Evolução da taxa de homicídios por idade - São Paulo

Evolução da taxa de homicídios por idade - Minas Gerais

ao Estado do Rio de Janeiro onde para todas as idades o comportamento da probabilidade parece indicar tendência decrescente.

O Estado de Minas Gerais apresenta freqüências de homicídio bastante inferiores às dos outros dois estados. Além disso, não parece existir um padrão comum de comportamento das freqüências não-controladas para as diversas idades ao longo do tempo.

A evolução das variáveis associadas aos indivíduos, como nível de escolaridade média e percentual de domicílios chefiados por mulheres apresenta comportamento linear crescente ao longo das duas décadas, e praticamente homogêneo entre as três regiões estudadas.¹³ No Brasil, a escolaridade média da PEA apresentou elevação aproximada de um ano de estudo a cada década. Em decorrência desse comportamento praticamente constante ao longo do tempo, essas variáveis apresentam correlação significativa com qualquer variável de tendência. O mesmo padrão de correlação é percebido para o coeficiente de Gini, que apresenta pequena variância ao longo do tempo se diferenciando apenas no nível entre as regiões. Esse padrão torna essa variável extremamente correlacionada com as *dummies* de região.

As demais variáveis econômicas apresentam comportamento cíclico ao longo das duas décadas. A taxa de desemprego aumenta significativamente

13 Os gráficos dessas variáveis estão apresentados no Apêndice 2.

com a recessão iniciada em 1981, caindo, porém, em 1982 e 1983, quando ainda se observa uma retração do produto. O Estado do Rio de Janeiro apresenta alterações significativamente menores do que São Paulo na taxa de desemprego em praticamente todo o período após 1986. Esse comportamento provavelmente está associado à elevada participação do setor informal na economia. O comportamento do rendimento médio da população ocupada é bastante similar entre os três estados, parecendo refletir os ciclos do produto da economia brasileira: o salário real tem uma redução forte de 1981 a 1984, recupera-se a partir de 1985 e se eleva com o cruzado; em seguida, torna-se estável, passando a decrescer a partir de 1990 até 1993, e novamente a partir de 1994, com o controle inflacionário, passa a apresentar tendência crescente. Todas as variáveis estão descritas no Apêndice 1.

3 - Metodologia

3.1- Modelo básico

O método de estimação utilizado neste trabalho é uma generalização do método qui-quadrado aplicado ao modelo Logit, também denominado Modelo de Berckson. Os modelos Logit são utilizados quando as variáveis dependentes são qualitativas, representadas por variáveis binárias — 1 se o evento ocorrer e 0, se não ocorrer. Nesse modelo, supõe-se que a probabilidade de ocorrência do evento depende das variáveis independentes segundo a seguinte forma funcional:

$$P_i = P(y_i = 1) = \Lambda(X' \beta) = \frac{e^{X' \beta}}{1 + e^{X' \beta}}$$

No caso de um modelo com múltiplas observações, podemos aplicar o modelo Logit com variáveis agrupadas. Nesse modelo, para cada idade temos n_i observações que correspondem ao total de homens residentes em cada região de determinada idade e o evento “morte por homicídio” ocorre para m_i observações segundo uma distribuição binomial.¹⁴ Seja p_i a probabilidade de ocorrência de m_i eventos mortes por homicídio em n_i observações e \hat{p}_i a frequência observada em uma dada realização. Como usual, podemos escrever:

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

¹⁴ O agrupamento realizado vai além do simples cálculo da frequência observada de homicídios porque a probabilidade foi estimada segundo as técnicas demográficas adequadas.

onde u é uma variável aleatória com média 0. Desse modo,

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

A única dificuldade com esse modelo é que, sob a hipótese de que as realizações da variável aleatória seguem uma distribuição binomial, a variância de u não é constante e o modelo é heterocedástico. Aplicando uma expansão de Taylor na função:

$$\text{Log} \frac{\hat{p}_i}{(1-\hat{p}_i)}$$

em torno de p_i , obtemos:

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

Assim, temos:

$$\text{Var}(u_i) = \left[\frac{p_i(1-p_i)}{n} \frac{1}{p_i^2(1-p_i)^2} \right] = \frac{1}{n_i p_i(1-p_i)}$$

que pode ser estimada por:

$$\text{Var}(u_i) = \frac{1}{n_i (\hat{p}_i (1-\hat{p}_i))}$$

A correção do modelo corresponde a uma estimação em mínimos quadrados em que se utiliza como fator de ponderação:

$$w = \left(\frac{1}{\text{Var}} \right)^{\frac{1}{2}}$$

Como no modelo básico estimado incluímos a probabilidade defasada, é necessário acrescentar um novo termo no cálculo da variância. A equação (3) deve ser reescrita da seguinte forma:

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

Com a inclusão da probabilidade defasada, o modelo passa a ser descrito por:¹⁵

$$F^{-1}(p_{it}) \equiv \beta X'_t + \gamma \left[\beta X'_{t-1} + \frac{1}{f_{t-1}(\hat{p}_{t-1})} \varepsilon_{t-1} \right] + \frac{\varepsilon_t}{f_t(\hat{p}_t)} \quad (5)$$

onde γ é o coeficiente da probabilidade defasada. Assim temos a seguinte equação para a variância:

$$\begin{aligned} \text{Var} \left[\gamma \frac{\varepsilon_{t-1}}{f_{t-1}(\hat{p}_{t-1})} + \frac{\varepsilon_t}{f_t(\hat{p}_t)} \right] &= \frac{\gamma^2}{f_{t-1}^2(\hat{p}_{t-1})} \text{Var}[\varepsilon_{t-1}] + \frac{1}{f_t^2(\hat{p}_t)} \text{Var}[\varepsilon_t] = \\ &= \frac{\gamma^2}{f_{t-1}^2(\hat{p}_{t-1})} \frac{\hat{p}_{t-1}(1-\hat{p}_{t-1})}{n_{t-1}} + \frac{1}{f_t^2(\hat{p}_t)} \frac{\hat{p}_t(1-\hat{p}_t)}{n_t} \end{aligned} \quad (6)$$

Como:

$$f^2(\hat{p}_t) = \hat{p}_t^2 (1-\hat{p}_t)^2$$

a fórmula final da variância é dada por:

$$\text{Var} = \frac{\gamma^2}{\hat{p}_{t-1}(1-\hat{p}_{t-1})n_{t-1}} + \frac{1}{\hat{p}_t(1-\hat{p}_t)n_t} \quad (7)$$

O modelo é estimado em dois estágios: estimamos a variância e no segundo estágio estimam-se os coeficientes das variáveis explicativas. Uma simples generalização do argumento utilizado por Amemya (1985, p. 276-277) mostra que a distribuição do estimador corrigido pela variância converge para uma distribuição normal.

3.2 - Testes de hipóteses

Neste trabalho utilizamos dois tipos de teste: especificação do modelo e robustez dos coeficientes, ambos baseados na mesma estatística — a soma do quadrado dos resíduos ponderados da regressão.

No teste de especificação classificamos os modelos segundo o ajuste das variáveis estimadas às variáveis observadas. Este teste segue o procedimento proposto por Li (1977). Denominando

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

¹⁵ Suprimimos o superescrito i para não abusar da notação.

onde \hat{p}_i é a frequência observada. Calcula-se a soma do quadrado dos desvios ponderada pela variância estimada do modelo, isto é:

$$\sum_{i=1}^T (\hat{L}_i - x'_i \beta_i) \sigma^{-2} (\hat{L}_i - x'_i \beta_i) = \chi^2$$

Em seguida, para cada modelo dividimos o valor da estatística qui-quadrado obtida pelo número de graus de liberdade, χ^2 / gl . O número de graus de liberdade corresponde a $T - K$, onde T é o número de células e K o número de parâmetros estimados. Para estabelecer a classificação dos modelos observamos os menores valores da estatística obtidos, que indicam um melhor ajustamento da especificação aos valores observados.

Para testar a significância das variáveis, implementamos o teste descrito em Amemya (1985). Neste caso, calculamos a soma do quadrado dos resíduos ponderada pela variância estimada do modelo completo para os dois modelos — completo e restrito — e verificamos se esta diferença é superior ao valor da estatística qui-quadrado. Isto é, definindo a soma do quadrado dos resíduos como:

$$SQR = \sum_{i=1}^T (\hat{L}_i - x'_i \beta_i) \sigma_i^{-2} (\hat{L}_i - x'_i \beta_i)$$

não podemos rejeitar a significância da variáveis restritas desde que:

$$SQR_{\text{modelo restrito}} - SQR_{\text{modelo completo}} > \chi_q^2$$

onde q é o número de variáveis restritas.

3.3 - Modelos estimados

Para cada base construída por idade foram rodadas 11 especificações diferentes incluindo as variáveis desemprego, desemprego defasado, salário real, coeficiente de Gini, *dummies* para os Estados do Rio de Janeiro e São Paulo, ano, índice de preços e probabilidade defasada. Em função dos problemas de correlação existentes entre escolaridade e taxa de domicílios chefiados por mulheres com as *dummies* de região e com a probabilidade defasada, essas variáveis não foram incluídas no modelo. Como mencionado anteriormente, essas variáveis apresentam comportamento crescente monotônico ao longo do tempo, tendo portanto correlação forte com as variáveis de tendência e nível incorporadas no modelo. O coeficiente de Gini, embora também apresente elevada correlação principalmente com as *dummies* de região, foi incluído em algumas especificações. O modelo com a taxa de desemprego passada foi incluído porque em algumas

economias a relação entre violência e taxa de desemprego passada é diferente da verificada para a taxa de desemprego contemporânea. A Tabela 1 sistematiza os 11 modelos estimados.

Tabela 1

Especificação dos modelos

| MODELO | VARIÁVEIS INCLUÍDAS |
|-----------|---|
| Modelo 1 | Salário real + desemprego + Gini |
| Modelo 2 | Salário real + desemprego + Gini + Prob (-1) |
| Modelo 3 | Salário real + desemprego + Gini + <i>dummyRio</i> + <i>dummySp</i> |
| Modelo 4 | Salário real + desemprego + Gini + <i>dummyRio</i> + <i>dummySp</i> + Prob(-1) |
| Modelo 5 | Salário real + desemprego + Gini + <i>dummyRio</i> + <i>dummySp</i> + + Prob(-1) + Ano |
| Modelo 6 | Salário real + desemprego + <i>dummyRio</i> + <i>dummySp</i> + Prob(-1) + Ano |
| Modelo 7 | Salário real + desemprego + <i>dummyRio</i> + <i>dummySp</i> + Prob(-1) |
| Modelo 8 | Salário real + desemprego + <i>dummyRio</i> + <i>dummySp</i> + Prob(-1) + INPC |
| Modelo 9 | Salário real + desemprego(-1) + <i>dummyRio</i> + <i>dummySp</i> + Prob(-1) |
| Modelo 10 | Salário real + desemprego + Gini + <i>dummyRio</i> + <i>dummySp</i> + + Prob(-1) + INPC |
| Modelo 11 | Salário real * [(100-Des/100)] + Gini + <i>dummyRio</i> + <i>dummySp</i> + + Prob(-1) + INPC |

Uma extensão natural dessas especificações é estimar um modelo contendo toda a base de dados, aproximando a forma funcional para a idade utilizando um polinômio de segundo grau. Desse modo, para a base contendo as 1.410 células foram regredidos dois modelos nos quais todas as variáveis independentes, exceto a probabilidade defasada, foram estimadas segundo uma especificação multiplicativa com a idade.

Foram regredidos os seguintes modelos:

Modelo 12

$$\begin{aligned}
 \text{Log}(\text{prob}/(1-\text{prob})) = & \alpha_1 \text{cte} + \alpha_2 \text{salreal} * \text{idade} + \alpha_3 \text{salreal} * \text{idade}^2 + \\
 & + \alpha_4 \text{desemprego} * \text{idade} + \alpha_5 \text{desemprego} * \text{idade}^2 + \alpha_6 \text{dummyRio} * \text{idade} + \\
 & + \alpha_7 \text{dummyRio} * \text{idade}^2 + \alpha_8 \text{dummySP} * \text{idade} + \alpha_9 \text{dummySP} * \text{idade}^2 + \\
 & + \alpha_{10} (\text{Log}(\text{prob}(-1)/(1-\text{prob}(-1)))) + \alpha_{11} \text{Gini} * \text{idade} + \alpha_{12} \text{Gini} * \text{idade}^2
 \end{aligned}$$

Modelo 13

$$\begin{aligned} \text{Log}(prob/(1-prob)) = & \alpha_1 cte + \alpha_2 salreal * idade + \alpha_3 salreal * idade^2 + \\ & + \alpha_4 desemprego * idade + \alpha_5 desemprego * idade^2 + \alpha_6 dummyRio * idade + \\ & + \alpha_7 dummyRio * idade^2 + \alpha_8 dummySP * idade + \alpha_9 dummySP * idade^2 \end{aligned}$$

A probabilidade defasada, calculada para cada idade específica, já incorpora o efeito da idade. A diferença entre os dois modelos é a introdução ou não do coeficiente de Gini.

4 - Resultados

Das 11 especificações propostas para a regressão considerando as bases por idade, somente os modelos 1 e 3 se revelaram como pouco ajustados às frequências observadas [ver Andrade e Lisboa (2000b)]. Estes modelos se distinguem dos demais por não incluírem a probabilidade defasada, que sem dúvida parece ser a variável com maior poder de explicação. A estatística qui-quadrado apresenta para os demais modelos baixa sensibilidade às alterações nas variáveis, não permitindo a identificação de uma especificação dominante. Esse comportamento provavelmente decorre da existência de multicolinearidade entre as variáveis *dummies*, Gini, ano e probabilidade defasada, confirmada pelas alterações dos coeficientes quando retiramos uma dessas variáveis da equação [ver Andrade e Lisboa (2000b)].

Surpreendeu-nos o elevado grau de ajustamento dos modelos e a similitude do comportamento das variáveis econômicas sugerindo a robustez dos modelos, que pode ser percebida pelo valor da estatística qui-quadrado. Este ajustamento é relativamente melhor nas caudas da distribuição etária.

Para analisar o comportamento dos coeficientes considerando as bases por idade, optamos por utilizar os modelos 7 e 10 como básicos.¹⁶ Os demais modelos, exceto as especificações 9 e 11, são variantes intermediárias desses dois, diferenciando-se, com relação ao modelo 10, pela exclusão de algumas variáveis lineares. O modelo 7 se distingue do modelo 10 pela exclusão das variáveis Gini e INPC. Como a variável Gini apresenta correlação com as *dummies* regionais, optamos por considerar o modelo 7 como a especificação de controle da análise.

No que se refere às variáveis econômicas, o principal resultado encontrado é a evidência de que a consideração da variável idade é extremamente importante no entendimento da relação entre violência e ciclos econômicos. O comportamento das variáveis econômicas, salário real e desemprego é bastante diferenciado tanto no sinal da relação quanto na intensidade [ver Andrade e Lis-

¹⁶ O teste de significância das variáveis foi feito apenas para estas duas especificações e estão apresentados em Andrade e Lisboa (2000b).

boa (2000b)]. Os resultados dos modelos regredidos evidenciam que para os homens mais jovens, 15 a 19 anos, as variáveis econômicas são importantes para explicar a violência. Para essas idades, o salário real apresenta o sinal esperado, negativo, sugerindo que o aumento do salário reduz os homicídios na população mais jovem.

Entre 20 e 29 anos, a importância do salário é praticamente nula. Após os 30 anos para algumas idades o coeficiente passa a ser positivo. Os testes de robustez sugerem que a variável salário é significativa para todas as bases no modelo de controle e significativa para os jovens de 15 a 17, 21, 23 e algumas idades acima dos 30 anos no modelo em que incluímos o coeficiente de Gini. Ou seja, a significância do salário real se altera quando introduzimos o coeficiente de Gini, sugerindo correlação entre essas variáveis.

Dessa forma, o aumento do salário real tem como efeito reduzir a taxa de homicídio entre os jovens. Para os mais velhos, porém, é possível que o aumento do salário real aumente a sua chance de ser vítima de homicídio. Como já discutimos, uma possível explicação para esse fenômeno é que os mais jovens tendem a ser vítimas de homicídio quando participam de atividades ilegais e essa participação aumenta com a queda do salário real. Os mais velhos, por outro lado, tendem a ser vítimas da violência com o aumento do retorno do crime: dada a parcela da população dedicada às atividades ilegais, o aumento do salário real aumenta o retorno de certas atividades ilegais e muitas das vítimas dessas atividades são homens acima de 30 anos.

Esses resultados são corroborados quando analisamos as especificações 12 e 13 que foram regredidas para a base completa aproximando o efeito da idade por meio de um polinômio de segundo grau [ver Andrade e Lisboa (2000b)]. O mesmo padrão de comportamento é observado para o salário, que para as idades mais jovens apresenta coeficiente negativo. No modelo completo, os coeficientes para o salário são significativos somente quando testamos as duas variáveis conjuntamente.

O desemprego apresenta comportamento bastante similar ao do salário nas especificações regredidas por idade, com coeficiente significativo e negativo para os jovens entre 15 e 20 anos. A partir dos 21 anos, o coeficiente do desemprego é praticamente zero e os testes de robustez sugerem que esta variável deixa de ser significativa tanto no modelo de controle quanto no modelo 10.

No modelo completo, o desemprego também apresenta coeficiente significativo e negativo, mas esse sinal permanece para idades mais avançadas. O resultado para as duas bases surpreende pela relação negativa obtida entre taxa de desemprego e probabilidade de morte, contrária ao esperado. Segundo o modelo de economia de crime de Becker (1968), a relação entre desemprego e ciclos econômicos apresenta dois efeitos contrários: por um lado, a redução do desemprego aumenta as oportunidades dos indivíduos no mercado de trabalho e portanto reduziria a probabilidade de mortes por homicídio, mas, por outro, aumenta o retorno esperado do crime tornando essas atividades mais atraentes.

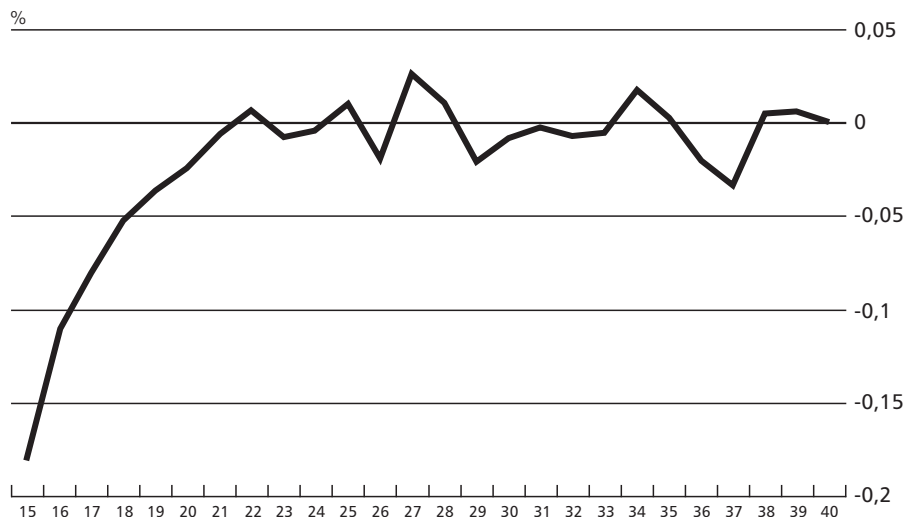
Entretanto, o último é um efeito de segunda ordem e não esperávamos que dominasse a relação.

Este mesmo resultado aparece em vários trabalhos em séries de tempo para a economia americana [Freeman (1994)] e diversos autores utilizam a taxa de desemprego defasada, em vez da taxa contemporânea. Os resultados encontrados nesse último caso, entretanto, são semelhantes aos obtidos anteriormente [ver modelo 9 em Andrade e Lisboa (2000b)]. Três fenômenos podem estar ocorrendo: *a*) como mencionamos, o efeito de segunda ordem da variação do desemprego sobre a taxa de homicídio pode de fato estar dominando o efeito de primeira ordem; *b*) a série utilizada pode ser reduzida, viesando os resultados, uma vez que temos apenas 17 anos, e em geral para a economia americana as séries de tempo têm em média 40 anos; e *c*) a elasticidade da taxa de desemprego pode ter se alterado ao longo da década, em decorrência, principalmente, do crescimento do setor informal.¹⁷ Estes resultados estão ilustrados nos Gráficos 8 a 11.

Os resultados encontrados para a probabilidade defasada confirmam a importância do efeito inércia para explicar a criminalidade principalmente a partir dos 20 anos. A probabilidade defasada é significativa e positiva para todas as

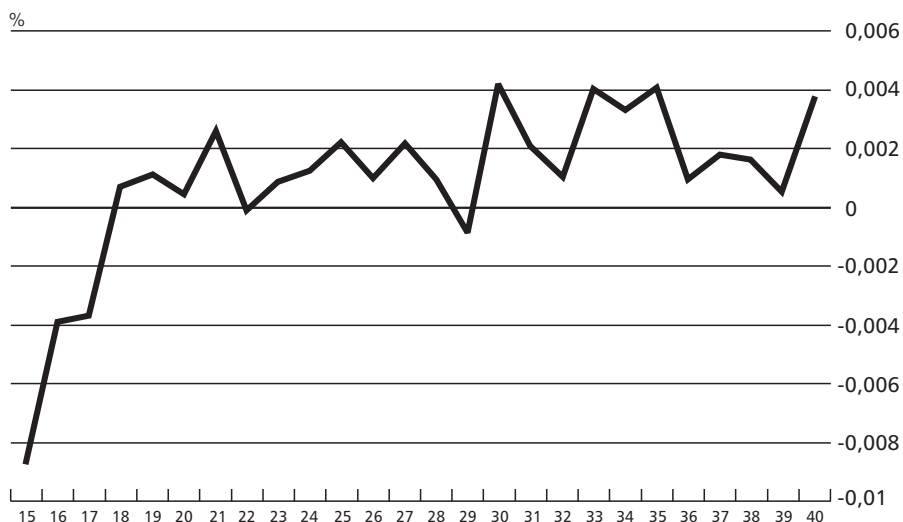
Gráfico 8

Coeficiente do desemprego por idade - modelo 7



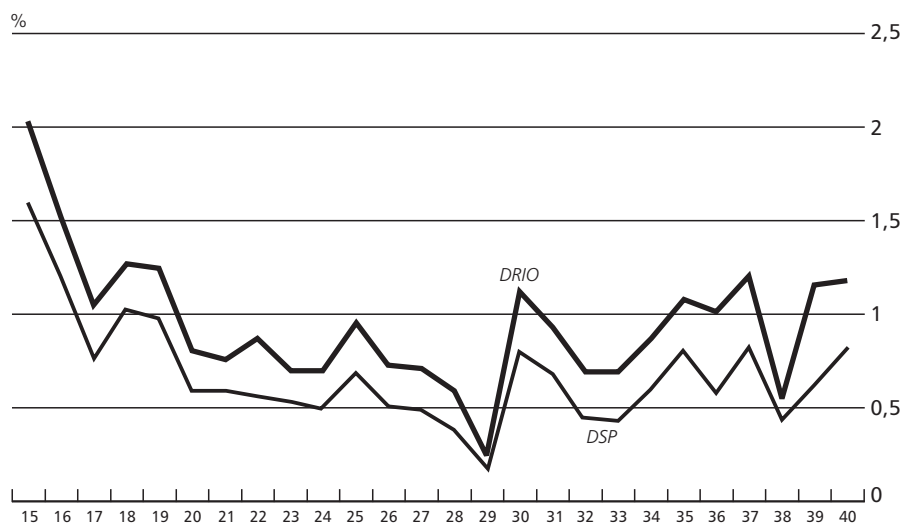
17 Uma forma de tentar controlar este problema é incluir no conjunto das variáveis independentes uma variável que controle o crescimento do setor informal. Tanto no Estado do Rio de Janeiro, como em São Paulo, a participação dos trabalhadores por conta própria no total da população ocupada é crescente nesse período, sugerindo um arrefecimento da taxa de desemprego nesses estados.

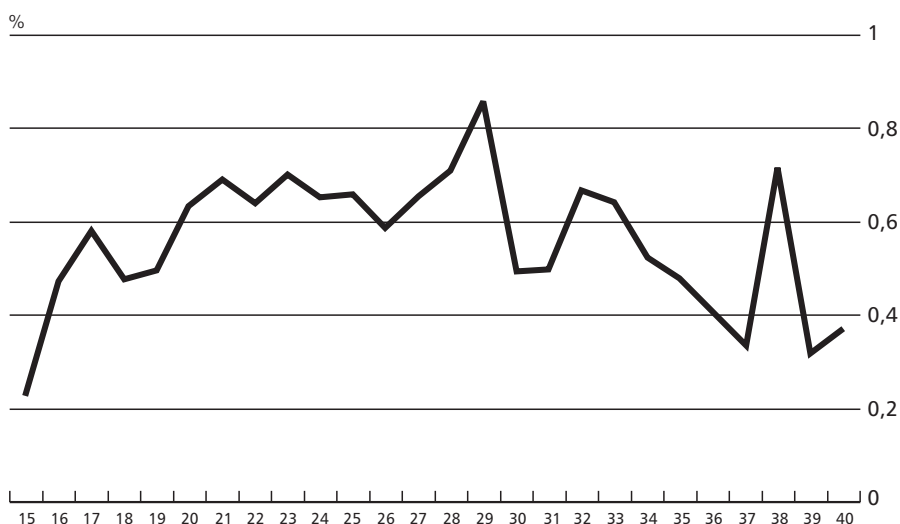
Gráfico 9

Coefficiente do salário real por idade - modelo 7

Desesperança de vida: homicídio em Minas Gerais, Rio de Janeiro e São Paulo no período 1981/97

Gráfico 10

Coefficiente das *dummies* RJ/SP por idade - modelo 7

Coefficiente da probabilidade defasada por idade - modelo 7

idades, apresentando, inicialmente, um impacto crescente com a idade. A partir dos 30 anos esse efeito volta a se reduzir. Tanto o teste de robustez como o teste de especificação revelam o poder explicativo da probabilidade defasada, que procura incorporar o efeito inércia: gerações que apresentam elevada taxa de homicídios quando jovem tendem a manter uma taxa maior por todo o seu ciclo de vida. Como as variáveis econômicas afetam a probabilidade de homicídios quando as gerações são jovens, e esse efeito é perpetuado por toda a geração por meio do efeito inércia, o impacto dessas variáveis é muito superior ao sugerido pela magnitude dos coeficientes.

Como discutimos na introdução, uma possível interpretação dos resultados relaciona a idade da geração à fração dedicada às atividades ilegais em cada ano. Suponhamos que seja custoso o retorno das pessoas dedicadas às atividades ilegais ao mercado legal de trabalho. Suponhamos, além disso, que a taxa de homicídio cresça com a fração da geração dedicada às atividades ilegais e que a mobilidade entre as atividades legais e ilegais se reduza com a idade: jovens deixam uma atividade legal por alguma ilegal com maior facilidade do que os mais velhos. Nesse caso, uma geração que quando jovem apresenta uma elevada fração dedicada à atividade ilegal, decorrente, por exemplo, de uma queda do salário real, tende a manter essa elevada fração ao longo de todo o ciclo de vida e, conseqüentemente, uma elevada taxa de homicídio ao longo desse mesmo ciclo.

Independente da validade dessa interpretação, porém, permanece a evidência de que uma geração que quando jovem apresente elevadas taxas de

homicídio tende a apresentar elevadas taxas durante todo o seu ciclo de vida, sugerindo a existência de uma persistência da violência específica a cada geração.

As *dummies* de região apresentam coeficiente positivo e significativo para todas as idades, mostrando que as regiões do Rio de Janeiro e São Paulo apresentam taxas de homicídio significativamente superiores às do Estado de Minas Gerais em praticamente todas as idades, mesmo quando controladas pelas variáveis econômicas utilizadas. Os coeficientes das *dummies* sugerem que esta diferença é mais importante para as idades mais jovens, 15 a 17 anos, e vai perdendo importância com o crescimento da idade. O mesmo padrão de comportamento é sugerido pelo modelo completo. Esse resultado é interessante porque o fenômeno da violência no Estado de Minas Gerais ainda se encontra bem menos disseminado que nos outros dois estados e, desse modo, os dados parecem sugerir, caso a nossa interpretação esteja correta, que os jovens em Minas Gerais ainda tentam ingressar inicialmente na atividade legal. A decisão de ir para a atividade do crime parece ser tomada em idades um pouco mais avançadas.

Surpreendeu-nos, no entanto, a significativa redução da diferença entre Rio de Janeiro e São Paulo quando os dados são controlados pela variáveis econômicas e pelo efeito inércia. Enquanto nos dados não-controlados a diferença oscila em torno de 100%, nos dados controlados essa diferença cai para cerca de 20%. Isso apesar de os dados não serem controlados por diferenças nas políticas de segurança, que talvez apresentem diferenças significativas em ambos os estados nesse período. Talvez essa diferença não seja tão significativa, ou ao menos seu impacto não o seja. Talvez, ainda, a política de segurança esteja correlacionada com as variáveis econômicas: estados mais ricos em certo período teriam maior acesso a instrumentos de segurança mais eficazes. O melhor entendimento dessas possibilidades requer, no entanto, o acesso a dados confiáveis sobre a política de segurança aos quais, no entanto, não tivemos acesso.

Tanto com a introdução da variável ano, como com a introdução do índice de preços, os coeficientes de todas as variáveis, exceto a constante, praticamente não se alteram [ver Andrade e Lisboa (2000b)]. A introdução do coeficiente de Gini no modelo também não altera de forma relevante os coeficientes das variáveis econômicas. O coeficiente que mais se modifica é o da *dummy* para o Estado de São Paulo, resultado esperado uma vez que essas variáveis apresentam correlação positiva. O coeficiente de Gini apresenta sinal positivo e significativo para todas as idades, entretanto, não é possível distinguir com clareza os efeitos da desigualdade de renda e estado específicos. O que se evidencia nos dados é que o Estado do Rio de Janeiro é mais violento que o Estado de São Paulo e os dois, por sua vez, mais violentos que o Estado de Minas Gerais. Essas diferenças, entretanto, podem não ser unicamente creditadas à desigualdade de renda. Para captar este efeito com mais segurança seria necessário expandir o número de regiões, de modo que a variabilidade da desigualdade de renda fosse ampliada.

5 - Relação com a literatura sobre economia do crime

Nos Estados Unidos uma vasta literatura procura estudar a relação entre variáveis econômicas e taxa de criminalidade. No Brasil, entretanto, a maior parte dos trabalhos concentra-se nas áreas de criminologia e aspectos sociais da criminalidade. Talvez esse aspecto da literatura brasileira deva-se, em parte, à pouca disponibilidade de base de dados confiáveis, principalmente no que se refere à política de segurança.

A maior parte dos trabalhos da literatura econômica americana tem como arcabouço básico a teoria dos incentivos, proposta inicialmente por Becker (1968) na qual a atividade do crime é considerada substituta da atividade legal.¹⁸ Nesse modelo, a cada instante os indivíduos tomam uma decisão entre alocar seu tempo na atividade legal ou alocá-lo à atividade do crime. A escolha sobre em qual atividade se engajar depende do retorno líquido esperado nas duas oportunidades. No cálculo do retorno líquido esperado do crime, os indivíduos consideram tanto os ganhos e perdas materiais quanto a probabilidade de ser preso e condenado por cometer atividade criminosa, e ainda os custos de reputação e acesso ao mercado de trabalho no longo prazo, caso sejam condenados.

Recentemente, alguns autores têm tentado incorporar a idéia de *inércia ou persistência* dos ciclos de violência. No modelo construído por Sah (1991) o resultado de persistência da criminalidade decorre da endogeneização da probabilidade de ser punido por ato criminoso. No arcabouço tradicional, a probabilidade de ser punido é exógena, constante ao longo do tempo e idêntica entre os indivíduos. Ao decidir se participam ou não da atividade do crime, os indivíduos consideram a sua percepção individual da probabilidade de punição, e não a probabilidade verdadeira. A *prior* dos indivíduos sobre a probabilidade de punição depende do comportamento das pessoas que eles têm capacidade de observar a cada período, enquanto a probabilidade verdadeira depende do nível de gastos realizados no sistema e da taxa de participação do crime. A cada período uma nova coorte passa a integrar o número total de pessoas ativas decidindo se participam ou não do crime. A taxa de participação no crime no período *T* depende de todas as coortes ativas. Nesse sentido, à medida que os agentes envelhecem, o conjunto de informação se eleva, e a *prior* dos indivíduos tende a se aproximar do valor da probabilidade real.

Sah supõe que para dado nível de gastos do sistema existe um nível máximo de apreensões pela polícia, de forma que o aumento da criminalidade reduz a probabilidade real de ser punido, assim como influencia a *prior* dos indivíduos e suas escolhas futuras. Desse modo, se a taxa de criminalidade passada for alta,

18 Tauchen e Witt (1994) e Grogger (1997) utilizam um arcabouço um pouco diferenciado. No modelo proposto, a atividade do crime não é um substituto da atividade legal. Os indivíduos podem realizar as duas atividades conjuntamente. Nesses tipos de modelo a escolha do indivíduo não é entre se engajar ou não na atividade do crime, mas qual o tempo ótimo de alocação nessa atividade.

a propensão do indivíduo ao crime no período corrente será mais elevada gerando o efeito de persistência da violência. O efeito das taxas de criminalidade passadas sobre a criminalidade corrente depende do tempo do ciclo ativo de cada coorte.¹⁹

Existem basicamente três tipos de abordagem empírica que buscam entender a relação entre os incentivos econômicos e a taxa de criminalidade: os trabalhos em séries temporais, os trabalhos em *cross-sections* e os trabalhos baseados no comportamento individual. Os resultados encontrados nesses três tipos de abordagem são bastante diferenciados. Freeman (1994) faz uma resenha dos principais resultados que são robustos para cada uma das três abordagens.

Nos trabalhos em séries de tempo, os resultados dependem do período em que foram estimados, da região e de como a taxa de criminalidade é mensurada, não fornecendo uma regra única entre as relações econômicas e a atividade do crime. Entretanto, essa parece ser a base de dados mais adequada para se explicar como em determinada região geográfica a taxa de criminalidade responde às oscilações econômicas.

Os trabalhos em *cross-sections* utilizam comumente dados de diversas localidades, sendo mais indicados para explicar a variância da taxa de criminalidade em diferentes regiões e não os seus determinantes. Nesse tipo de abordagem, os resultados são bastante robustos e a taxa de criminalidade apresenta comportamento anticíclico. A relação entre desigualdade de renda e criminalidade também parece ser robusta nesse tipo de análise. Fajnzylber, Lederman e Loayza (1999) a partir de uma base de dados para 45 países encontram uma relação positiva entre níveis de desigualdade de renda, mensurados pelo coeficiente de Gini, e taxa de homicídio intencional. A fragilidade desse tipo de análise é que os resultados podem estar, em alguma medida, associados a diferentes características da população não-controladas nos experimentos gerando, portanto, correlação espúria entre as variáveis analisadas.

Os modelos com dados individuais são os que apresentam relação mais forte entre as variáveis econômicas e a decisão de participar do crime. Essas pesquisas, em geral realizadas em penitenciárias, não consideram todo o universo dos indivíduos se restringindo aos que já se decidiram pela atividade do crime.

O elevado número de trabalhos reflete tanto a gravidade do problema da criminalidade, sobretudo na economia americana, como também a dificuldade em lidar com o tema sob uma abordagem econômica. Entre as principais dificuldades, citamos: a) a não-existência de uma medida dos retornos do crime que permita separar os efeitos de redução das oportunidades legais e os efeitos

19 Outro argumento da literatura que pode ser alçado para explicar a persistência dos ciclos de violência é o da interação social proposto por Glaeser, Sacerdote e Scheinkman (1996). Estes autores estavam interessados em explicar a elevada variância das taxas de criminalidade ao longo do tempo e entre regiões não explicada pela variância das condições econômicas. Segundo eles, existe uma covariância positiva entre as decisões dos agentes em participar do crime e assim a variância do crime é um múltiplo da variância se os agentes tomassem as decisões de forma independente.

de elevação da probabilidade de ser punido e de aumento dos retornos do crime; b) a não-existência de medidas exatas de criminalidade dado o elevado número de sub-registros; e c) a dificuldade em se mensurar a probabilidade de ser punido, já que as variáveis comumente utilizadas — gastos da polícia, número de policiais *per capita* e taxa de prisões — podem estar correlacionadas com o aumento da violência ou mesmo apresentar erros de mensuração. Um fato estilizado da literatura é a associação entre maiores retornos do crime e economias onde existe maior produção e comercialização de drogas. Em função da proliferação do uso de drogas, alguns autores [Freeman (1994)] defendem que os retornos ao crime têm aumentado nas últimas duas décadas.

Os resultados de três trabalhos recentes para a economia americana vão ao encontro das evidências encontradas neste artigo. Grogger (1997) mostra que o comportamento dos homens jovens apresenta elevado grau de resposta aos incentivos econômicos e que, portanto, a queda do salário real nas duas últimas décadas é determinante importante da elevação da taxa de criminalidade. Tauchen e Witt (1994), por outro lado, demonstram, que os jovens que alocam maior tempo no trabalho ou na escola apresentam menor probabilidade de participar do crime. Ambos os estudos acompanharam durante o período de sete anos uma coorte de homens nascida em 1945 na Pensilvânia. Freeman (1996) defende que a taxa de criminalidade tem aumentado em decorrência da redução das oportunidades de trabalho para os homens jovens menos qualificados. Segundo o autor, o retorno das atividades legais para os indivíduos menos qualificados se reduziu em contrapartida à elevação dos retornos do crime. Além disso, a elasticidade da oferta de trabalho às variações no salário é suficientemente elevada para aumentar a propensão à atividade do crime. Acrescente-se ainda o argumento de que a atividade do crime pode não ser necessariamente substituta da atividade legal, podendo ser exercida conjuntamente pelos jovens para garantir elevação da renda.

Sobre o caso brasileiro, os resultados ainda parecem incertos. Carneiro e Phebo (1999) encontram resultados bastante distintos dos esperados. Os autores analisam a relação entre taxa de criminalidade, medida como homicídios por 100 mil habitantes nos municípios das regiões metropolitanas do Rio de Janeiro e São Paulo, e flutuações econômicas, utilizando um modelo em painel desbalanceado para quatro anos, estimado por mínimos quadrados ponderados. Os resultados indicam relação negativa entre desigualdade de renda e criminalidade e comportamento pró-cíclico da criminalidade.

Beato e Reis (1999) analisam a correlação entre taxa de criminalidade, medida de forma desagregada como crimes contra o patrimônio e contra a pessoa, e desenvolvimento econômico nos municípios da região metropolitana de Belo Horizonte em um modelo *cross-section*.²⁰ A correlação encontrada depende da

20 Os autores utilizam como *proxies*, para grau de desenvolvimento, o índice de desenvolvimento humano, a taxa de analfabetismo e a taxa de mortalidade infantil.

forma como a taxa de criminalidade é mensurada. O trabalho, no entanto, não procura controlar simultaneamente todas as variáveis independentes, examinando apenas a correlação e a dispersão entre cada par de variáveis.

Uma provável dificuldade comum aos dois trabalhos é a unidade de análise. A integração do mercado de trabalho nas regiões metropolitanas provavelmente impede que as flutuações econômicas sejam diferenciadas ao nível dos municípios que pode resultar em uma correlação espúria, como discutimos na introdução. Suponhamos que dois municípios próximos apresentem taxas de criminalidade significativamente distintas. Ora, o município com maior criminalidade, na ausência de migração, deve apresentar alguma compensação de bem-estar para os seus moradores, por exemplo maior salário real ou menor taxa de desemprego. Dessa forma, uma base de dados em *cross-section* que indique a existência de correlação positiva entre salário real e criminalidade não significa que um crescimento do salário real aumente a criminalidade. De fato, talvez exatamente o oposto seja verdadeiro: caso o salário real aumente em ambos os municípios, a criminalidade pode cair sem contradizer a correlação obtida em *cross-section* que, inclusive, pode ser preservada para o novo salário real. O resultado obtido talvez apenas signifique que, se um município apresenta maior criminalidade, o bem-estar dos seus habitantes é compensado pelo maior salário real. Por essa razão, parece-nos importante considerar uma base de dados com série de tempo em que se observam flutuações concomitantes das variáveis econômicas em diversos períodos em todas as regiões.

Além disso, como discutimos na introdução, a utilização como base de dados das taxas de homicídios por 100 mil habitantes pode viesar os resultados, em decorrência tanto da existência de composições de sexo e idade distintas nas regiões quanto da omissão da variável idade. Por essa razão, neste trabalho construímos a base de dados calculando a frequência de morte por homicídio a cada idade em cada ano e região.

A construção dessa base de dados nos permitiu considerar dois aspectos que parecem extremamente relevantes no entendimento dos ciclos de violência: incorporar o efeito inércia e estimar coeficientes específicos a cada idade. Ao incorporar o efeito inércia como específico à geração, a inércia dos ciclos de violência, já observada em outros trabalhos, ganha nova interpretação. O resultado mais robusto encontrado neste artigo se refere, precisamente, à importância da probabilidade defasada para explicar os ciclos de violência, obtida a partir da construção das frequências de homicídio por coortes. Os coeficientes estimados para a probabilidade defasada sugerem que os indivíduos tomam a decisão de entrar no crime dos 15 aos 20 anos e, uma vez que decidam participar dessa atividade, tendem a continuar na mesma. Foge do escopo deste trabalho explicar os determinantes dessa inércia.

A estimação do modelo para cada idade nos permite sugerir algumas relações entre os ciclos econômicos e os ciclos de violência. Os resultados encontrados evidenciam que as variáveis econômicas são relevantes para explicar a

criminalidade até no máximo 20 anos de idade. Este resultado corrobora os trabalhos de Grogger e Tauchen e Witte e parece ser extremamente importante do ponto de vista da intervenção pública. Os dados sugerem que políticas voltadas para a população masculina jovem podem reduzir a criminalidade no longo prazo. Entretanto, como o impacto dessas políticas é defasado ao longo do ciclo de vida, seus efeitos somente podem ser percebidos no longo prazo, o que talvez torne sua implementação menos atraente para os governantes.

Os resultados encontrados para a relação entre desigualdade de renda e violência, embora apresentem o sinal positivo e significativo, não são muito robustos. A enorme disparidade observada nos níveis de violência dos três estados pode, talvez em parte, ser explicada pela desigualdade de renda, mas este resultado pode estar associado a características estado-específicas relacionadas à política de segurança. Como já enfatizamos anteriormente, porém, a ausência de base de dados confiáveis inviabilizou uma extensão do trabalho para investigar essa relação.

6 - Agenda de pesquisa

Muito embora os resultados encontrados neste trabalho nos pareçam robustos, a precariedade dos dados econômicos por coorte e a falta de acesso a dados sobre a política de segurança enfraquecem nossa confiança nos modelos testados. É possível apontar pelo menos quatro extensões deste trabalho:

a) Construção de uma base de dados em que as variáveis econômicas reflitam as características de cada idade. Neste artigo as variáveis econômicas foram aproximadas pela média.

b) Inclusão de variáveis instrumentais para a política de segurança.

c) Ampliação do número de unidades da Federação analisadas, de forma a se obter maior dispersão e maior desigualdade de renda, e se testar medidas alternativas de desigualdade.

d) Inclusão de uma variável que controle a participação do setor informal na absorção de mão-de-obra.

Em princípio, as variáveis relevantes para a entrada em atividades ilegais devem ser condicionadas às características individuais e não à média da população em cada região e ano. Uma desagregação interessante seria condicionar os dados de frequências de homicídios à escolaridade e às características dos pais, assim como condicionar as variáveis econômicas à idade e à escolaridade. Infelizmente, porém, os dados disponibilizados pelo SIM apresentam problemas precisamente no preenchimento desses dados. Uma solução intermediária seria condicionar ao menos as variáveis econômicas por idade, além de região e ano.

No que se refere à política de segurança, algumas variáveis alternativas, como efetivo policial, poderiam ser tentadas como *proxy* do grau de comprometimento do Estado com políticas punitivas. A política de segurança, porém, não

se esgota na punição, e nem mesmo fica claro em que medida o efetivo policial descreve adequadamente esse comprometimento. Nesse aspecto, cabe uma investigação mais cuidadosa de critérios e indicadores mais adequados da política de segurança.

A inclusão de uma variável de controle para a participação do setor informal na absorção de mão-de-obra nos permitiria investigar com mais cuidado o sinal da relação entre desemprego e taxas de homicídio, já que a evolução da taxa de desemprego no Brasil parece sugerir uma mudança de sensibilidade nos ciclos de retração do produto ao longo dos anos 80 e 90. O maior volume de pessoas no setor informal, em sua maior parte, realizando atividades temporárias, faz com que esse contingente deixe efetivamente de procurar emprego, passando a ser considerado como parcela da população ocupada.

Por fim, no que se refere ao peso da desigualdade na taxa de homicídio duas extensões parecem adequadas. Por um lado, aumentar o número de unidades da Federação analisadas de modo a permitir diferenciar a desigualdade de outros aspectos regionais e, por outro, testar índices alternativos de desigualdade. A principal dificuldade com essa extensão está no elevado número de subenumeração de óbitos em diversos estados brasileiros, sobretudo nos do Norte e Nordeste.

Apêndice 1

Descrição das variáveis

Variáveis econômicas

- Taxa de desemprego
Taxa de desemprego aberto em cada região metropolitana calculada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) através dos dados obtidos na Pesquisa Mensal de Emprego (PME).
Conceito: pessoas que efetivamente procuraram emprego nos últimos 30 dias e não exerceram nenhum trabalho nos últimos sete dias.
- Rendimento
Índice do rendimento médio real no trabalho principal por região metropolitana a da população ocupada. Deflator utilizado: Índice Nacional de Preços ao Consumidor (INPC). Base do índice: jul./94=100. Dados disponíveis a partir do ano de 1982.
Fonte: IBGE/PME.
Dados coletados da revista *Conjuntura Econômica* da Fundação Getulio Vargas – dezembro de 1997 e novembro de 1998.
- Índice Nacional de Preços ao Consumidor
Variação mensal acumulada anualmente para o Brasil.
Fonte: IBGE.
Dados coletados da revista *Conjuntura Econômica* da Fundação Getulio Vargas (vários anos).
- Coeficiente de Gini
Índice calculado para as unidades da Federação baseado em dados da Pesquisa Nacional por Amostra de Domicílios do (PNAD) do IBGE.
Para os anos de 1991 e 1994 o índice foi interpolado linearmente.
Dados fornecidos pela equipe de pesquisa de Ricardo Paes de Barros (IPEA/RJ).

Variáveis de família

- Escolaridade
Nível médio de anos de estudos médio da PEA das unidades da Federação calculado com base na PNAD.
Para os anos de 1991, 1994 e 1997 as informações foram interpoladas linearmente.
- Taxa de Domicílios com Chefes do Sexo Feminino
Número de domicílios cuja pessoa de referência é do sexo feminino/número de domicílios permanentes total.
Fonte: PNAD/IBGE (1981/97). Para os anos de 1991 e 1994 os dados foram interpolados linearmente.

- Probabilidade de morte

Os dados de probabilidade de morte foram calculados segundo as técnicas usuais de demografia utilizando a base de dados sobre mortalidade disponibilizada (SIM) pela Fundação Nacional de Saúde através do Datasus.²¹ Para os estados de São Paulo e Minas Gerais foram utilizados apenas os registros de morte por homicídio e no Estado do Rio de Janeiro foram utilizados os registros de mortes por homicídio e mortes por outras violências. Estes grupos correspondem aos seguintes códigos: homicídios e lesões provocadas intencionalmente: CID9 9600-9699 e na CID 10 X850-Y099; outras violências: CID9 9700-9999 e na CID10 Y100-Y369.

O elevado número de homicídios classificados como outras violências no Estado do Rio de Janeiro parece ser um problema com as bases de dados de atestado de óbito neste estado. Para contornar essa dificuldade padronizamos as mortes por acidente no Estado do Rio de Janeiro, incluídas no grupo outras violências, de acordo com a distribuição do Estado de São Paulo. A categoria outras violências inclui mortes ocasionadas em sete circunstâncias: lesões por intervenção legal, lesões de guerra, envenenamentos, enforcamentos, lesões devido a queda, lesões de intencionalidade ignorada causadas com armas de fogo e armas brancas e lesões por meios não especificados. Os dois primeiros grupos apresentam registro praticamente nulo nos dois estados. O grupo dos envenenamentos, enforcamentos e lesões devido a queda constitui o das mortes causadas por acidente. São precisamente os dois últimos subitens que apresentam a maior participação na probabilidade de morte associada a esta categoria no Rio de Janeiro em quase todos os anos. Carneiro e Phebo (1998) encontram o mesmo problema ao utilizar essa base de dados. Eles supõem que todas as mortes provocadas por armas de fogo teriam sido homicídios, as mortes provocadas por objetos cortantes teriam sido acidentes e 50% das demais causas cuja intencionalidade é ignorada seriam homicídios. A comparação da série temporal do Rio de Janeiro com a série de São Paulo revela o (relativo), reduzido volume de registros de morte na categoria outras violências em São Paulo, onde seu peso na probabilidade de morte raramente ultrapassa 1% ao longo de todo o período. Esse resultado sugere a existência de um problema de mensuração estado-específica. Como os índices de homicídio em São Paulo são relativamente altos e estáveis, não sugerindo uma eventual subenumeração maior neste caso do

21 Como estimamos a probabilidade de morte para a população acima dos 10 anos, podemos utilizar a fórmula de cálculo básica da construção da tábua de vida, isto é:

$$q_x = \frac{D_x}{pop_x + 0.5 \times D_x}$$

Onde q_x é a probabilidade de morrer em determinada idade específica, D_x corresponde ao número de mortes em determinado ano e em determinada idade e pop_x corresponde à população residente naquele ano em determinada idade.

que no Rio de Janeiro, o problema parece se encontrar no Estado do Rio de Janeiro. Fizemos, então, a seguinte simulação. Supusemos que a participação dos homicídios no total das causas *homicídios e outras violências* no Rio de Janeiro fosse, em cada ano, para cada idade e sexo, idêntica à obtida em São Paulo. Nossa hipótese é de que as mortes acidentais correspondem a um evento aleatório de igual distribuição no Rio e em São Paulo, sendo, porém, correlacionadas com o nível de homicídios. Tendo por base essa hipótese, estimamos a probabilidade de morte por homicídio para os três estados segundo as técnicas demográficas usuais. Os dados de população residente foram obtidos por meio de uma expansão log-linear dos dados de população para cada idade dos anos censitários, isto é, 1980 e 1991, e da Contagem Populacional de 1996.

Para os Estados do Rio de Janeiro, São Paulo e Minas Gerais, a subenumeração de óbitos contabilizada pelo IBGE é relativamente pequena e a perda de informação decorrente da base de dados utilizada alcança em média 7% do universo global.

Apêndice 2

Gráfico 1

Chefe de família - São Paulo/Rio de Janeiro/Minas Gerais

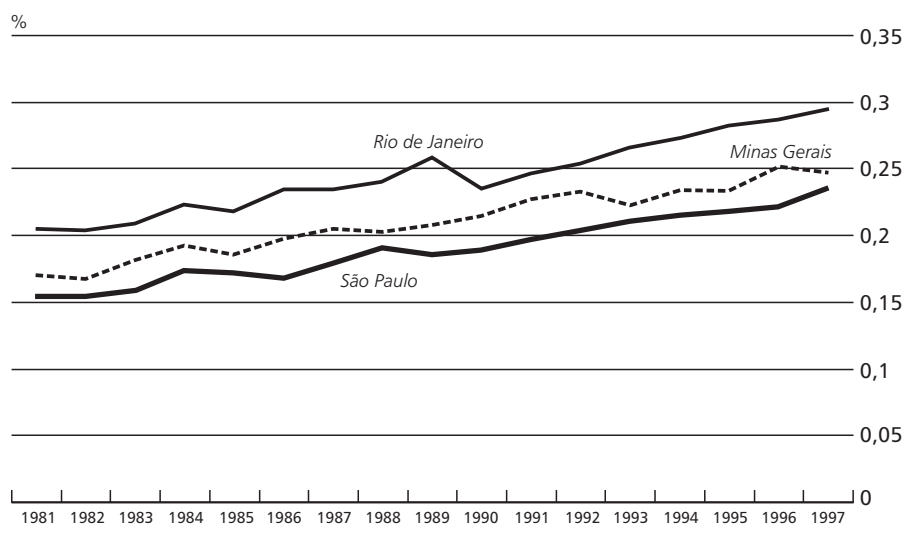
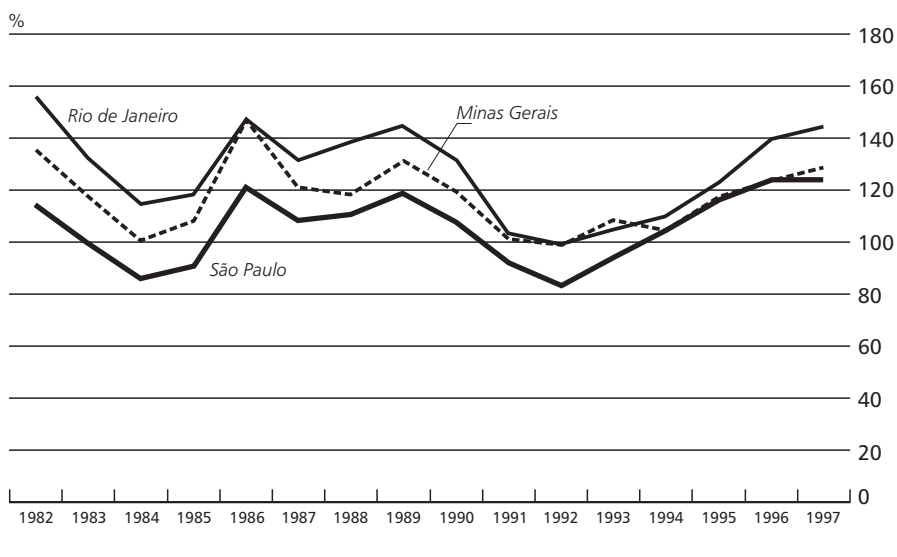


Gráfico 2

Rendimento médio - São Paulo/Rio de Janeiro/Minas Gerais

Desesperança de vida: homicídio em Minas Gerais, Rio de Janeiro e São Paulo no período 1981/97

Gráfico 3

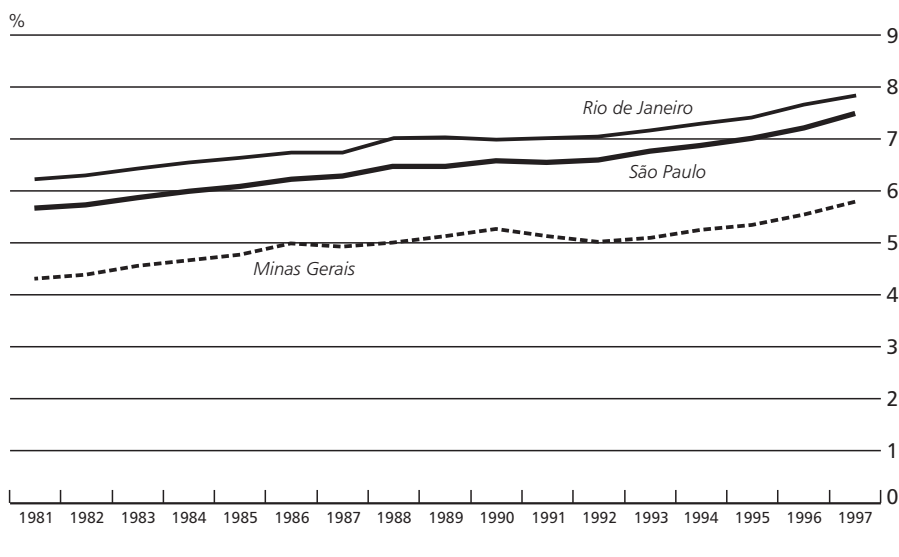
Escolaridade - São Paulo/Rio de Janeiro/Minas Gerais

Gráfico 4

Desigualdade e pobreza no Brasil

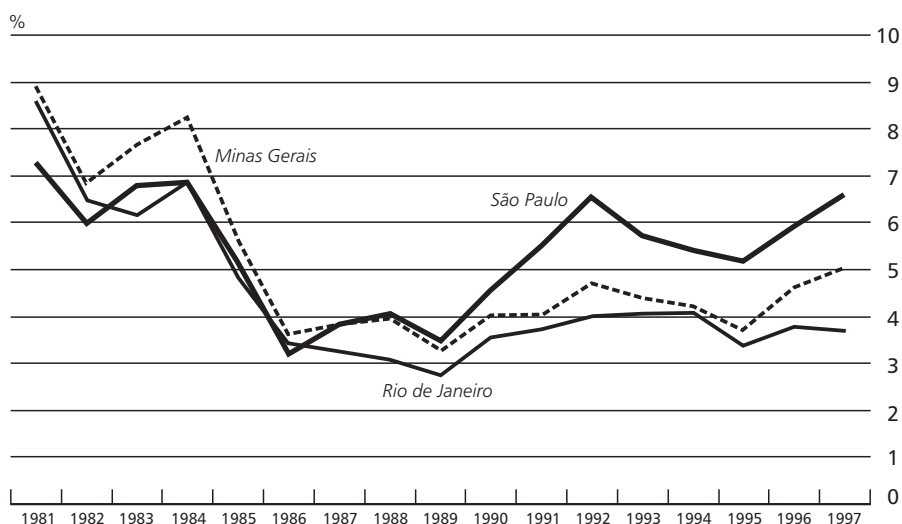
Taxa de desemprego - São Paulo/Rio de Janeiro/Minas Gerais

Gráfico 5

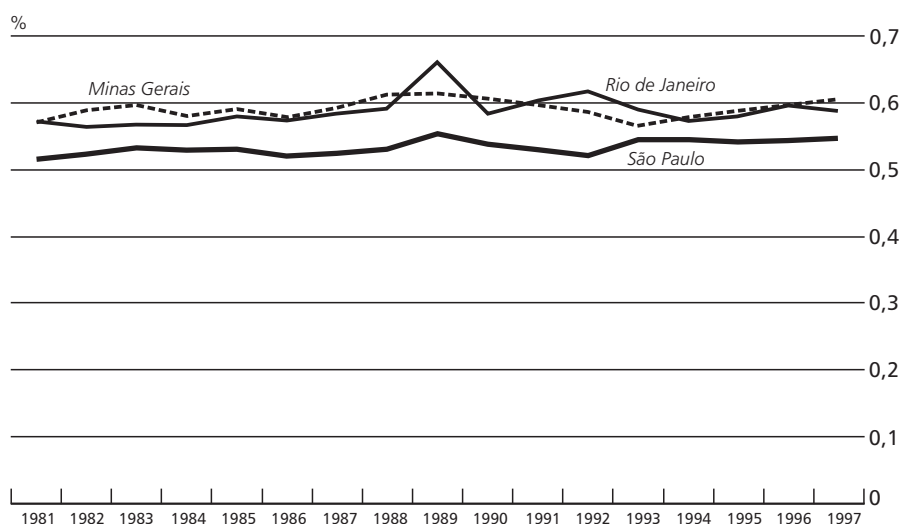
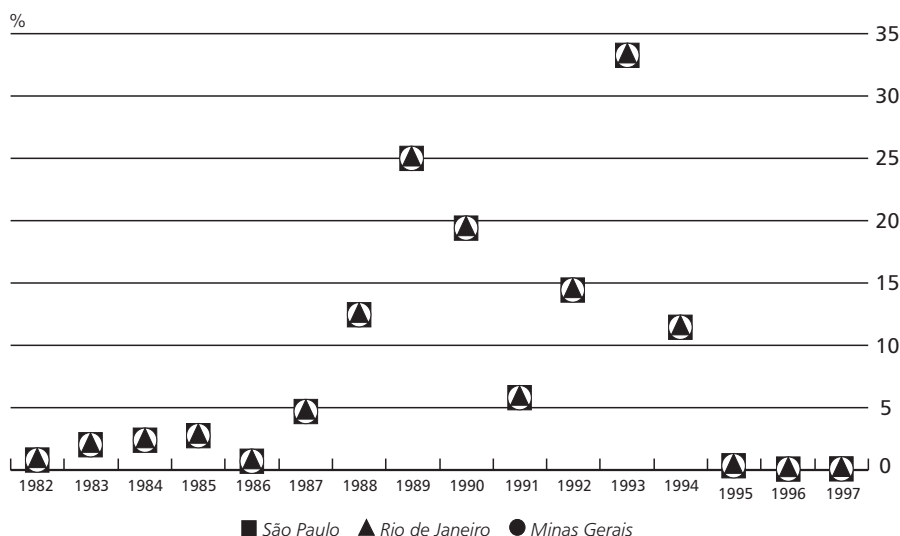
Índice de Gini - São Paulo/Rio de Janeiro/Minas Gerais

Gráfico 6

INPC - São Paulo/Rio de Janeiro/Minas Gerais**Bibliografia**

AMEMYA, T. *Advanced econometrics*. Oxford, Basil Blackwell, 1985.

ANDRADE, M. V., LISBOA, M. B. *Alguns fatos estilizados sobre o setor saúde no Brasil*. EPGE/FGV, 2000a, mimeo.

———. Desesperança de vida: homicídio em Minas Gerais, Rio de Janeiro e São Paulo no período 1981/97. Rio de Janeiro. *Ensaio Econômico*. Rio de Janeiro: EPGE/FGV, 2000b.

BEATO, C., REIS, I. A. *Desigualdade, desenvolvimento socioeconômico e crime*. Texto apresentado no seminário Desigualdade e Pobreza no Brasil, promovido pelo IPEA/RJ, Rio de Janeiro, ago. 1999.

BECKER, G. Crime and punishment: an economic approach. *Journal of Political Economy*, v. 76, n. 2, p.169-217, 1968.

CARNEIRO, L. P., PHEBO, L. *Magnitude, custos econômicos e políticas de controle da violência no Rio de Janeiro*. Banco Interamericano de Desenvolvimento, 1999 (Série Documentos de Trabalho, 347).

FAJNZYLBER, P., LEDERMAN, D., LOAYZA, N. *Inequality and violent crime*. 1999, mimeo.

———. *Determinants of crime rates in Latin America and the world*. World Bank, Latin American and Caribbean Studies (Viewpoint Series), 1998.

- FREEMAN, R. B. *Crime and the job market*. Cambridge: National Bureau of Economic Research, Mar. 1994 (Working Paper, 4.910).
- . *Why do so many young American men commit crimes and what might we do about it?* Cambridge: National Bureau of Economic Research, Feb. 1996 (Working Paper, 5.451).
- FUNDAÇÃO SEADE: *Anuário Estatístico do Estado de São Paulo*, 1993.
- GLAESER, E., SACERDOTE, B., SCHEINKMAN, J. Crime and social interactions. *Quarterly Journal of Economics*, v. 111, n. 2, 1996.
- GOVERNO DO ESTADO DO RIO DE JANEIRO. *Anuário Estatístico do Estado do Rio de Janeiro*. Secretaria de Estado e Planejamento, 1997.
- GROGGER, J. *Market wages and youth crime*. Cambridge: National Bureau of Economic Research, Mar. 1997 (Working Paper, 5.893).
- IBGE. *Censo Demográfico, 1980 e 1991*.
- . *Contagem Populacional*. CD-ROM, 1986.
- LI, M. A logit model of homeownership. *Econometrica*, v. 45, n. 5, 1977.
- MADDALA, G. S. *Limited dependent and qualitative variables econometrics*. New York: Cambridge University Press, 1983.
- MINISTÉRIO DA PREVIDÊNCIA E ASSISTÊNCIA SOCIAL. *Nona Classificação Internacional de Doenças*, 1990.
- MINISTÉRIO DA SAÚDE. *Sistema de Informação sobre Mortalidade 1979 – 1997. Dados de Declaração de óbito*. CD-ROM, 1997.
- RYDER, N. B. The cohort as a concept in the study of social change. *American Sociological Review*, v. 30, n. 6, Dec. 1965.
- SAH, R. K. Social osmosis and patterns of crime. *Journal of Political Economy*, v. 99, n. 6, 1991.
- SHRYOCK, A. S., SIEGEL, J. S. *The methods and materials of demography*. Washington D.C.: U.S. Bureau of Census, Second Printing, 1973.
- STENFENSMEIER, D., STREIFEL, C., SHIHADDEH, E. S. Cohort size and arrest rates over the life course: the easterlin hypothesis reconsidered. *American Sociological Review*, v. 57, p. 306-314, June 1992.
- TAUCHEN, H., WITT, A. *Work and crime: an exploration using panel data*. Cambridge: National Bureau of Economic Research, 1994 (Working Paper, 4.794).
- WITT, R., WITT, A. D. *Crime, imprisonment, and female labor force participation: a time-series approach*. Cambridge, nov. 1998 (NBER Working Paper, 6.786).