

Uma estimativa dos determinantes da taxa de criminalidade brasileira: uma aplicação em painel dinâmico

Leandro Kume*

EPGE/FGV

Resumo

A partir de dados de painel para estados brasileiros no período 1984-1998, estimamos os determinantes da taxa de criminalidade brasileira utilizando o Método Generalizado dos Momentos em Sistema proposto por Blundell e Bond (1998). Desta forma, foi possível amenizar problemas econométricos, que afetam a maioria dos trabalhos nesta área, como endogeneidade das variáveis explicativas e erros de medida gerados pela subnotificação das taxas de crimes. As estimativas obtidas permitem concluir que o grau de desigualdade de renda e a taxa de criminalidade do período anterior geram um efeito positivo sobre a taxa de criminalidade do período presente, enquanto que o PIB per-capita, o nível de escolaridade, o grau de urbanização e o crescimento do PIB têm efeitos negativos. Além disso, a aplicação do Método Generalizado dos Momentos em Primeira Diferença (Arellano e Bond, 1991), nesse caso, não revelou ser apropriado, devido a presença de instrumentos fracos, apresentando um estimador para a taxa de criminalidade defasada bem abaixo do estimador Within-Groups. O uso de instrumentos adicionais em primeira diferença para as equações em níveis através do MGM-SIS mostrou ser mais informativo.

Abstract

Using a panel data for Brazilian states in the period 1984-1998, we estimate the determinants of Brazilian crime rate. The estimates was derived by System Generalized Method of Moments propose by Blundell and Bond (1998). In this way, it was possible to avoid econometric problems, that affect the majority of the works in this area, such as endogeneity and measure of error generated by the subnotification on crime rates. The results show that inequality and crime rates of the previous period have positive effect on crime rate of the present period, while the per capita GDP, the level of schooling, the degree of urbanization and the GDP growth have a negative effect. Moreover, the use of the First Difference Generalized Method of Moments estimator (Arellano and Bond, 1991), in this case, did not reveal to be appropriated due to weak presence of weak instruments. The coefficient was below the Within-Groups estimator. The use of System MGM show to be more informative in terms of bias and efficiency.

Classificação do JEL: O10; K42; C23.

Palavras chave: Crime, punição, MGM em sistema.

Keywords: Crime, punishment, System GMM.

Classificação da Anpec: Economia do Trabalho, Economia Social e Demografia (Área 6).

*Endereço: *Praia de Botafogo*, 190, sala 1124, *Rio de Janeiro* - RJ, Brasil, CEP: 22.250-900. Email: *kume@fgvmail.br*.

1 Introdução

Desde o trabalho inicial de Becker (1968), a literatura econômica tem analisado a criminalidade como sendo uma atividade como outra qualquer, onde o ato de se cometer um delito é baseado numa decisão racional entre custos e benefícios. Um indivíduo, portanto, se tornaria um criminoso se tal atividade lhe proporcionasse um retorno financeiro relativamente maior, levando-se em conta a probabilidade de ser preso e a severidade da punição.

Os trabalhos empíricos que analisam as causas da criminalidade têm encontrado resultados contraditórios devido a diversos obstáculos. Primeiro, dados sobre criminalidade sofrem do problema de subnotificação que varia de acordo com o tipo e o local do delito. Segundo, é provável que, além de ser afetada, a criminalidade também influi em uma série de variáveis econômicas. Por exemplo, a escolha das empresas em investir em determinadas regiões pode ser influenciada pelas taxas de criminalidade, assim como a decisão de morar em cidade pequena ou grande. Logo, devemos considerar a possibilidade das variáveis que determinam as taxas de criminalidade sejam endógenas. Por fim, a falta de controle das variáveis omitidas que variam pouco ao longo do tempo, mas que são correlacionadas com outras variáveis explicativas, pode também viesar as estimativas.

Basicamente, os fatores que influenciam a criminalidade podem ser divididos em dois grupos: os relacionados à severidade e a eficácia da punição e os vinculados aos benefícios e custos de oportunidade da prática de crime. Enquanto que Ehrlich (1973) e Mathieson e Passell (1976) estimaram uma elasticidade do crime em relação ao valor esperado da punição de, respectivamente, $-0,5$ e $-0,3$ corroborando o modelo teórico de Becker, Archer e Gartner (1984) não encontraram resultado significativo ao estudarem o efeito da pena de morte sobre a taxa de homicídio.

Um problema econométrico grave que afeta a maioria das estimações empíricas sobre a relação criminalidade e punição é a simultaneidade dessas duas variáveis. A presença de altas taxas de criminalidade, por exemplo, também induzem os governantes a contratar mais policiais ou a aumentar a severidade da punição. Logo, o uso do método de Mínimos Quadrados Ordinários (MQO) para estimar o efeito da punição sobre a criminalidade fornecerá estimadores viesados. Propondo instrumentos para o tamanho da população carcerária, o número de policiais per-capita e as taxas de condenação, Levitt (1996, 1997, 1998) isolou seus efeitos exógenos e encontrou resultados empíricos condizentes com o trabalho de Becker (1968).

Os estudos ligando os benefícios e custos de oportunidade à criminalidade também têm apresentado alguns resultados contraditórios. Ehrlich (1975), por exemplo, apresentou um resultado positivo entre anos de estudo de uma população adulta e crimes contra propriedade enquanto que Fajnzylber, Lederman e Loayaza (2000), ao contrário, revelaram que este efeito não é significativo sobre os índices de homicídios. Analisando o efeito das taxas de desemprego e das desigualdades de renda, Fleisher (1966) e Ehrlich (1973) mostraram uma redução da criminalidade decorrente de uma queda de tais variáveis, embora o impacto do nível de renda não tenha sido significativo.

No Brasil, a literatura econômica sobre a criminalidade ainda é incipiente. Dentre os trabalhos existentes, dois merecem destaques. Através de dados de série de tempo para os estados de Rio de Janeiro e São Paulo, Cerqueira e Lobão (2003) apresentaram evidências de que o aumento da desigualdade de renda tem um impacto positivo sobre a criminalidade enquanto que, para o nível de renda da

população e gastos em segurança pública no estado, o efeito é negativo. Araújo e Fajnzylber (2001), utilizando o Método Generalizado dos Momentos em Primeira Diferença (MGM-DIF), sugerido por Arellano e Bond (1991) para controlar problemas de endogeneidade, estimaram um painel dinâmico para estados brasileiros e mostraram que a renda per-capita é positivamente correlacionada com a atividade crimininosa. Além disso, seus resultados indicam a presença de um componente inercial na taxa de criminalidade¹. Uma variação de 1% da taxa de criminalidade do período anterior tem um impacto positivo de 0,25% na mesma taxa no período presente. Contudo, de acordo com Blundell e Bond (1998), os estimadores provenientes de MGM-DIF, sob certas circunstâncias, podem apresentar propriedades fracas em termos de viés e de eficiência. Como forma de amenizar este problema, os autores sugerem uso do Método Generalizado dos Momentos em Sistema (MGM-SIS) que utiliza condições de momentos adicionais.

O objetivo deste trabalho é analisar os determinantes da taxa de criminalidade brasileira, com base em um painel dinâmico por estados no período 1984-998, aplicando o MGM-SIS. Com a aplicação dessa metodologia, esse estudo se diferencia em relação às estimativas efetuadas no Brasil, por amenizar problemas frequentemente encontrados nas estimações como endogeneidade das variáveis explicativas, omissão de variáveis relativamente constantes ao longo do tempo e erros de medida na variável dependente. Dessa forma, será construído, na próxima seção, um simples modelo teórico sobre o comportamento do indivíduo frente ao crime de forma a explicitar as variáveis independentes a serem utilizadas no teste empírico. A seguir, uma breve resenha da metodologia econométrica será apresentada com o objetivo de destacar as vantagens do estimador MGM-SYS sobre MGM-DIF. Na terceira e quarta parte do trabalho, serão descritos, respectivamente, os dados utilizados e os resultados das estimações comparando-se os dois métodos. Por fim, serão apresentadas as principais conclusões.

2 Decisão individual de se cometer um crime

Assumindo que o indivíduo i é neutro ao risco e racional, a utilidade esperada associada a um crime (EU_i) pode ser expresso como o payoff esperado obtido do butim (l) vezes a probabilidade de não ser preso ($1 - pr$); menos o custo de execução e planejamento do crime (c), o custo de oportunidade (w), o valor esperado da punição no caso de ser preso ($pr * pu$) e o valor moral de tal ato.

$$EU_i = (1 - pr) * l - c - w - pr * pu - m$$

Se o benefício líquido de tal atividade for positiva ($EU_i > 0$) então o crime será cometido.

Existem diversas variáveis sócio-econômicas que afetam esse modelo, sendo que algumas são conhecidos e outras são ambíguas. O passado criminal de um indivíduo (c_{t-1}), por exemplo, influencia positivamente a decisão de se cometer um crime de três formas. Primeira, um indivíduo com ficha criminal tende a ser discriminado no mercado de trabalho, o que implica em menor custo de oportunidades (w menor). Segundo, um criminoso de longa data, assim como em qualquer outro emprego, acaba se especializando, diminuindo assim o custo de executar e planejar um crime (menor c),

¹Fajnzylber, Lederman e Loayza (2000 e 2002) encontraram resultados semelhantes utilizando dados internacionais.

bem como a probabilidade de ser preso (menor pr). Por fim, uma pessoa que já cometeu seu primeiro crime terá naturalmente um valor moral menor para se engajar em um segundo delito (menor m). A taxa de criminalidade agregada do período anterior (C_{t-1}) também influencia a decisão individual de se cometer um crime no presente, na medida que, transmite uma sensação de impunidade (menor pr). Dessa forma, a criminalidade tem um componente inercial de sinal positivo.

O crescimento e o nível da atividade econômica são variáveis que têm um efeito ambíguo na decisão de se cometer um crime, pois uma variação positiva cria maiores oportunidades de emprego no setor formal (maior w), mas também aumenta o butim esperado (maior l).

Já o efeito da desigualdade de renda sobre os índices de criminalidade depende do status social relativo de cada um, mas em termos agregados será positivo. Para uma pessoa rica, um aumento da desigualdade de renda onde vive, pouco importará na decisão de se cometer um crime, enquanto que, para um pobre, representará uma perda salarial (menor w). O aumento da desigualdade de renda também terá um efeito negativo nos valores morais dos mais pobres (menor m), pois se aprofunda a sensação de que o rico está se beneficiando do trabalho dos pobres.

O nível educacional do indivíduo é outra variável que, a princípio, tem um efeito ambíguo sobre o crime. Primeiro, amplia o valor moral de se cometer um crime (maior m). Segundo, cria condições para se obter maiores oportunidades de emprego (maior w). Terceiro, diminui o custo de se cometer um crime (menor c). Quarto, aumenta o lucro do crime (maior l). Quinto, reduz a probabilidade de ser preso (menor pr).

A força do aparato policial e da justiça em coibir a criminalidade pode aumentar a probabilidade do criminoso ser preso (maior pr), bem como uma punição mais severa (maior pu), diminuindo os incentivos do indivíduo em se engajar numa atividade ilegal. Em muitos países, existe um apelo da população pela adoção da pena capital como forma de se inibir a criminalidade.

Como os ambientes de maior aglomeração de pessoas facilitam a fuga e dificultam a identificação do criminoso (menor pr), altas taxa de urbanização estariam ligadas a elevados níveis de criminalidade. Além disso, a interação entre criminosos e futuros criminosos seria maior em áreas urbanas (menor c)².

Desta forma, enquanto que variáveis como o nível educacional, a força do aparato policial e da justiça teriam um efeito negativo sobre a criminalidade, o impacto da desigualdade de renda, do nível de crimes passados e da taxa de urbanização seria positivo. Já o efeito da atividade econômica seria ambíguo.

Dado que os dados disponíveis são agregados, não é possível estimar o efeito das variáveis econômicas sobre a decisão individual de se cometer um crime. Dessa forma, o seguinte modelo, obtido do somatório da função de se cometer um crime de cada indivíduo, será utilizado nas estimações:

$$crime_{it} = \alpha_0 + \alpha_1 crime_{i,t-1} + \beta X_{it} + \eta_i + \mu_t + v_{it} \quad (1)$$

onde $crime_{it}$ é a taxa de criminalidade no local i no período t , X_{it} é a matriz representando as demais variáveis explicativas, η_i é o efeito fixo ao longo do tempo, μ_t é o efeito do tempo e v_{it} é o termo de erro.

²Ver Glaeser, Sacerdote e Scheinkman (1996).

3 Metodologia

3.1 MGM em primeira diferença

Nesta seção será feita uma breve discussão das vantagens em se utilizar o método de MGM-SIS para estimar um painel dinâmico. Suponha o seguinte modelo:

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it} \text{ para } |\alpha| < 1, i = 1, \dots, N \text{ e } t = 2, \dots, T, \quad (2)$$

$$E[\eta_i] = E[v_{it}] = E[\eta_i v_{it}] = 0, \text{ para } i = 1, \dots, N \text{ e } t = 2, \dots, T. \quad (3)$$

onde y_{it} é a taxa de criminalidade "verdadeira" no estado i e ano t , η_i é o efeito fixo de cada estado sobre a criminalidade e v_{it} é o resíduo. Além disso, suponha ainda que o erro v_{it} não é autocorrelacionado temporalmente,

$$E[v_{it}v_{is}] = 0 \text{ para } i = 1, \dots, N \text{ e } s \neq t, \quad (4)$$

e a condição inicial, y_{i1} , é dada por

$$E[y_{i1}v_{it}] = 0 \text{ para } i = 1, \dots, N \text{ e } t = 3, \dots, T. \quad (5)$$

Note que o método de mínimos quadrados, neste caso, forneceria estimadores viesados devido a correlação existente entre $y_{i,t-1}$ e v_{it-1} . A partir das condições (2) a (5), Arellano e Bond (1991) propõem a utilização de variáveis com defasagens em pelo menos dois períodos, $(t-2)$, como instrumento para as equações (2) em primeira diferença. Para uma amostra de tamanho T , existem $m = 0,5(T-1)(T-2)$ condições de momentos suficientes para identificar e estimar o parâmetro α :

$$E[y_{i,t-s}\Delta v_{it}] = 0 \text{ para } t = 3, \dots, T \text{ e } s \geq 2 \quad (6)$$

Reescrevendo na forma matricial temos:

$$E[Z_i' \Delta v_i] = 0 \quad (7)$$

onde Z_i é uma matriz de dimensão $(T-2) \times m$ dado por

$$Z_i = \begin{bmatrix} y_{i1} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & y_{i1} & y_{i2} & \dots & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & y_{i1} & \dots & y_{i,T-2} \end{bmatrix} \quad (8)$$

e $\Delta v_i = (\Delta v_{i3}, \Delta v_{i4}, \dots, \Delta v_{iT})'$ é um vetor $(T-2) \times 1$. O estimador MGM-DIF, portanto, minimizará a norma de $\Delta v' Z W Z \Delta v$ onde W é a matriz de ponderação, Z é uma matriz $M \times N(T-2)$ definida por $[Z_1', Z_2', \dots, Z_N']$ e Δv é a matriz de erros $N(T-2) \times 1$ dado por $[\Delta v_1', \Delta v_2', \dots, \Delta v_N']$. Além disso, esse estimador será consistente para $N \rightarrow \infty$ e T fixo.

Contudo, para amostras finitas, Blundell e Bond (1998) apresentaram evidências que o estimador MGM-DIF teria propriedades fracas em termos de viés e precisão decorrente da baixa correlação entre os instrumentos propostos por Arellano e Bond (1991) e as variáveis explicativas. Em um modelo AR(1), por exemplo, isso ocorreria devido a existência do parâmetro autoregressivo (α) tendendo a unidade ou de uma variância do efeito fixo (η_i) crescendo a uma taxa relativamente maior que a variância do choque (v_{it})³. Através de simulações de Monte Carlo, Blundell e Bond (1998) também mostraram que o estimador MGM-DIF seria negativamente viesado nos casos onde o número de períodos de tempo fosse pequeno⁴.

Para se identificar a magnitude deste tipo de viés sobre os resultados, a literatura sugere utilizar outros métodos de estimação para efeito de comparação. Em um modelo AR(1), utilizar MQO em nível superestimaria α devido a presença do efeito fixo (Hsiao, 1986) enquanto o estimador Within Groups, ao contrário, para períodos de tempo relativamente pequeno, seria negativamente viesado (Nickell, 1981). Essa propriedade desses dois estimadores é válida também na presença de outras variáveis explicativas embora necessite que tais variáveis não sejam correlacionadas com η_i e sejam estritamente exógenas com respeito a v_{it} . Contudo, Blundell e Bond (1998b), Bond, Hoeffler e Temple (2001) e outros sugerem seu uso para efeitos comparativos. Espera-se, portanto, que um estimador consistente para α deveria se localizar num intervalo entre os estimadores de MQO em nível e de With Goups (WG). Deste modo, se o estimador MGM-DIF estivesse próximo ou abaixo do estimador WG então a qualidade dos instrumentos usados deveria ser analisada.

³Suponha que $T = 3$. Neste caso, as condições de momento, apontados em (6), se reduziriam a uma única condição de identificação para α . Podemos então expressar a seguinte equação na forma reduzida do estimador de variáveis instrumentais:

$$\Delta y_{i2} = \pi y_{i1} + r_i \text{ para } i = 1, \dots, N$$

$\Delta y_{i2} = \pi y_{i1} + r_i$ para $i = 1, \dots, N$ Note que subtraindo y_{i1} de (2) obtemos a seguinte equação:

$$\Delta y_{i2} = (\alpha - 1)y_{i1} + \eta_i + v_{i2} \text{ para } i = 1, \dots, N$$

Para um α suficiente grande ou uma variância de η_i que aumente a uma taxa relativamente maior que a variância de v_{it} , o estimador de MQO para o coeficiente π se aproxima de zero. O instrumento y_{i1} , deste modo, será fracamente correlacionado com Δy_{i2} . Assumindo estacionariedade e supondo $\sigma_\eta^2 = \text{var}(\eta_i)$ e $\sigma_v^2 = \text{var}(v_{it})$, obtemos

$$p \lim \hat{\pi} = p \lim (\widehat{\alpha - 1}) = (\alpha - 1) \frac{k}{\left(\frac{\sigma_\eta^2}{\sigma_v^2}\right) + k} \text{ com } k = \frac{(1 - \alpha)^2}{(1 - \alpha^2)}$$

Observe que $p \lim \hat{\pi} \rightarrow 0$ se $\alpha \rightarrow 1$ ou $\left(\frac{\sigma_\eta^2}{\sigma_v^2}\right) \rightarrow \infty$. Logo, sob certas condições, os instrumentos obtidos a partir da (6) são fracos

⁴Para $T = 4$, $N = 100$ e o verdadeiro valor de α igual a 0.9, por exemplo, o estimador MGM em primeira diferença, para 1000 simulações de Monte Carlo, obteve uma média de 0,23 (com desvio padrão igual a 0,83) mostrado na Tabela 2 de Blundell e Bond (1991).

3.2 MGM em sistema

Como forma de melhorar a performance do estimador MGM-DIF para amostras finitas, Blundell e Bond (1998) sugere uma hipótese adicional:

$$E [\Delta y_{i2} \eta_i] = 0 \text{ para } i = 1, \dots, N. \quad (9)$$

Para que (9) seja válida, basta que a média de y_{it} seja constante para cada indivíduo ao longo do tempo. A primeira vista, esta condição pode parecer forte, tendo em vista que a maioria dos modelos empíricos envolve variáveis não estacionárias como PIB, investimento, entre outras. Contudo, a adoção de dummies de tempo corrige este problema transformando as variáveis em desvios em relação às médias dos i locais de para cada período t . Note ainda que esta condição de estacionariedade sobre a média não impõe nenhuma restrição sobre o segundo momento de y_{i1} ⁵. Embora y_{it} seja correlacionado com o efeito fixo, η_i , dados as hipóteses do modelo, a correlação entre Δy_{it} e η_i é igual zero. Juntando esta hipótese ao modelo apresentado em (2) a (5), obtemos $T - 2$ condições de momentos adicionais:

$$E [\Delta y_{i,t-1} v_{it}] = 0 \text{ para } i = 1, \dots, N \text{ e } t = 3, 4, \dots, T. \quad (10)$$

Desta forma, variáveis defasadas em primeira diferença podem ser usadas como instrumentos para as equações em níveis. Utilizando tanto (10) como (6) obtemos a seguinte matriz de instrumentos:

$$Z_i^+ = \begin{bmatrix} Z_i & 0 & 0 & \dots & 0 \\ 0 & \Delta y_{i2} & 0 & \dots & 0 \\ 0 & 0 & \Delta y_{i3} & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \Delta y_{i,T-1} \end{bmatrix} \quad (11)$$

Onde Z_i é dado por (8). Reescrevendo na forma matricial a seguinte condição de momento:

$$E [Z_i^+ v_i^+] = 0 \quad (12)$$

onde $v_i^+ = (\Delta v_{i3}, \dots, \Delta v_{iT}, v_{i3}, \dots, v_{iT})'$. O estimador MGM gerado a partir deste conjunto de instrumentos é chamado na literatura por MGM-SIS⁶. Através de simulações de Monte Carlo, Blundell e Bond (1998) e Blundell, Bond e Windmeijer (2000) apresentaram evidências que esse estimador, para amostras finitas, teria uma performance melhor que o estimador MGM-DIF tanto em termos de vies como de eficiência. Esse ganho seria ainda maior conforme α se aproximasse da unidade.

⁵Ver Blundell e Bond (1998) e Bond, Hoeffler e Temple (2001).

⁶Dado que o conjunto total Z_i é usado como instrumentos para as equações em primeira diferença, para que não haja redundância, somente a defasagem mais recente dos instrumentos em níveis é aproveitada para as equações em níveis (Arellano e Bover, 1995).

3.3 Incluindo variáveis explicativas adicionais

Como extensão ao modelo anterior, adicionamos outra variável explicativa, x_{it} , além da variável dependente defasada.

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \eta_i + v_{it} \text{ para } |\alpha| < 1, i = 1, \dots, N \text{ e } t = 2, \dots, T, \quad (13)$$

onde x_{it} é correlacionada com η_i e é endógena no seguinte sentido de

$$E[x_{it}v_{is}] \neq 0 \text{ para } i = 1, \dots, N \text{ e } s \leq t. \quad (14)$$

Além disso, v_{it} segue a estrutura apresentado em (3) a (4).

Efetuando a primeira diferença a fim de se eliminar o efeito fixo, η_i , obtemos, além de (6), a condição de momento

$$E[x_{i,t-s}\Delta v_{it}] = 0 \text{ para } t = 3, \dots, T \text{ e } s \geq 2. \quad (15)$$

Ou seja, a variável endógena x_{it} defasada também pode ser utilizada como instrumento para para as variáveis em primeira diferença.

Analogamente a condição (9), vamos supor que a correlação entre x_{it} em primeira diferença e o efeito fixo para cada i seja igual a zero.

$$E[\eta_i\Delta x_{it}] = 0 \quad (16)$$

Então, a seguinte condição de momento também estará disponível:

$$E[\Delta x_{i,t-1}v_{it}] = 0 \text{ para } i = 1, \dots, N \text{ e } t = 3, \dots, T. \quad (17)$$

Portanto, a partir das condições de momentos disponíveis em (6), (10), (15) e (17) podemos obter novamente um estimador MGM-SIS consistente para os parâmetros do modelo.

3.4 Erro de medida na variável dependente

A subnotificação presente nos dados de criminalidade requer certos cuidados nas estimações. De fato, Soares (2002) apresentou evidências de uma correlação positiva entre o grau de reportagem de crimes e o nível de desenvolvimento dos países. Uma vez controlado este problema, a renda per capita não revelou um efeito significativo sobre as taxas de criminalidade. Como forma de amenizar esta questão, primeiramente, optou-se pelo uso de uma taxa de criminalidade cujo índice de subnotificação é menor, segundo, foi utilizado um tratamento econométrico de forma a lidar com o problema.

Supondo que o erro de medida em y_{it} é função de um choque aleatório i.i.d, δ_{it} , e um efeito fixo de cada estado, ψ_i , temos:

$$y_{it}^* = y_{it} + \psi_i + \delta_{it} \quad (18)$$

Substituindo em (2) temos:

$$y_{it}^* = \alpha y_{i,t-1}^* + \beta x_{it} + \mu_i + \xi_{it} \quad (19)$$

onde $\mu_i = \eta_i + (1 - \alpha)\psi_i$ e $\xi_{it} = \delta_{it} - \alpha\delta_{i,t-1} + v_{it}$. Note que o erro agora sofre de correlação serial devido a presença de δ_{it} e $\delta_{i,t-1}$. Logo, $y_{i,t-2}^*$ não é um instrumento válido para a equação em diferença. O mesmo vale para o caso onde existe o mesmo tipo de erro de medida para x_{it} . Analogamente às seções anteriores, temos agora as seguintes condições de momentos que serão exploradas pelo estimador GMM-SIS:

$$E [y_{i,t-s}^* \Delta \xi_{it}] = 0 \text{ para } t = 4, \dots, T \text{ e } s \geq 3, \quad (20)$$

$$E [x_{i,t-s} \Delta \xi_{it}] = 0 \text{ para } t = 3, \dots, T \text{ e } s \geq 3, \quad (21)$$

$$E [\Delta y_{i,t-2}^* \xi_{it}] = 0 \text{ para } i = 1, \dots, N \text{ e } t = 4, \dots, T, \text{ e} \quad (22)$$

$$E [\Delta x_{i,t-2} \xi_{it}] = 0 \text{ para } i = 1, \dots, N \text{ e } t = 4, \dots, T. \quad (23)$$

3.5 Testes de especificação

Por fim, como forma de se testar a robustez do modelo, Arellano e Bond (1991) sugerem dois tipos de testes. O teste de Sargan é utilizado com o objetivo de verificar a validade dos instrumentos. A falha em rejeitar a hipótese nula indicará que os instrumentos são robustos. Além disso, como supõe-se, inicialmente, que o erro v_{it} não seja autocorrelacionado, é feito um teste de correlação serial de primeira ordem e outro de segunda ordem sobre os resíduos em primeira diferença, Δv_{it} . Espera-se que os erros em primeira diferença sejam correlacionados em primeira ordem e não autocorrelacionados em segunda ordem.

4 Dados

Uma das principais dificuldades nesta área de pesquisa é a falta de dados confiáveis coletados de forma sistemática. No Brasil este problema é agravado pela ausência de um órgão federal que coordene e divulgue informações referentes a criminalidade. Além disso, como cada estado é responsável por sua própria estrutura de segurança pública, mudanças metodológicas e falta de homogeneidade das séries referentes a crimes são frequentes, impedindo a comparação direta dos dados divulgados por cada uma dessas instituições. Utilizar dados de criminalidade, com já mencionado, requer certos cuidados adicionais devido a presença de subnotificação cujo grau varia de acordo com o tipo de delito e a confiança no sistema judicial. De modo geral, crimes violentos como homicídios tende a ter uma taxa de reportagem maior que assalto ou furto.

Como forma de contornar esses obstáculos foram coletados, junto ao Sistema de Informação de Saúde (SIM) do Ministério da Saúde, homicídios intencionais por estado entre 1984 e 1998. Essa base de dados permite medir a criminalidade do país com um grau maior de homogeneidade e confiabilidade em relação aos óbitos dos registros de ocorrência da polícia de cada estado. Junto a estimativa populacional fornecida pelo IBGE, foi possível a construção de uma taxa homicídios por 100 mil habitantes (*lnrazhom*). As demais variáveis foram obtidas do IPEA-Data com exceção da taxa de urbanização que foi adquirida do IBGE. Como medida para educação e desigualdade de renda foram utilizadas respectivamente o número médio de anos de estudo para população a partir dos 25 anos (*estudo*) e a razão entre a renda média apropriada pelos 20% mais ricos e os 20% mais pobres (*desig*) para os anos 1984-1990, 1992-1993 e 1995-1998. O logaritmo do PIB real per capita (*lnpibp*) e da taxa de crescimento do PIB real anual (*crescpib*), medidos a preço de 2000, foram utilizados como proxy para o nível de atividade econômica e emprego. Pela falta de disponibilidade do número de policiais por estado, optou-se pelo uso dos gastos anuais per capita em segurança pública de cada estado deflacionado pelo IGP-DI (*seg*) fornecido pela Fundação Getúlio Vargas. Por fim, a taxa de urbanização anual por estado (*urb*) foi obtida pela interpolarização linear dos dados do censo demográfico de 1980, 1991 e 2000.

Desta forma, temos um painel desbalanceado de 1984 a 1998 para 26 estados⁷. Na tabela 1 são apresentados algumas estatísticas descritiva dos dados. O logaritmo da taxa de homicídio por 100 mil habitantes apresentou uma média de 20,2 enquanto que seu desvio padrão entre estados foi praticamente o dobro do desvio padrão ao longo do tempo. Isso levanta suspeitas de que a criminalidade é influenciada por características de cada estado.

⁷Como existe pouca informação sobre como os dados foram agregados a partir de 1989 para o Estado de Goiás com a criação do Estado de Tocantins, optou-se por excluir os dois estados da amostra a partir dessa data.

Tabela 1 - Estatísticas descritivas

Variável		Média	Std. Dev.	Min.	Max.	Obs
lnrazhom	Total	20.26	0.68	18.93	21.83	TxN=354
	between		0.61	19.37	21.53	N=26
	within		0.31	19.28	22.15	T médio=136,1
raz	Total	22.04	5.78	8.99	44.55	TxN=304
	between		4.24	15.21	33.39	N=26
	within		3.95	12.08	35.05	T médio=116,9
lnpibp	Total	10.14	0.68	8.75	11.64	TxN=354
	between		0.61	9.09	11.45	N=26
	within		0.31	9.11	12.19	T médio=136,1
estudo	Total	4.71	1.26	2.06	8.10	TxN=304
	between		1.13	2.80	7.45	N=26
	within		0.56	3.22	6.69	T médio=116,6
urb	Total	71.70	12.51	34.72	96.21	TxN=354
	between		12.00	43.51	95.51	N=26
	within		3.72	60.24	84.67	T médio=136,1
crescpib	Total	3.81	8.99	-17.88	45.23	TxN=328
	between		1.94	-0.88	9.28	N=26
	within		8.77	-18.96	39.76	T médior=126,1
lnseg	Total	11.90	0.74	7.87	13.67	TxN=354
	between		0.56	11.01	13.25	N=26
	within		0.49	8.39	14.02	T médior=136,1

Fonte: DATA-SUS. Elaboração própria

5 Resultados

Nas estimações feitas por MGM, optou-se pelo MGM em um estágio com correção para heterocedasticidade⁸. Embora teoricamente o estimador em dois estágios seja assintoticamente mais eficiente, simulações de Monte Carlo apontam para um pequeno ganho em performance, além de uma taxa de convergência relativamente mais lenta para sua distribuição assintótica. Além disso, para amostras finitas, o desvio-padrão do estimador em dois estágios pode apresentar um viés fortemente negativo⁹.

Na Tabela 2 são apresentados os principais resultados das regressões do modelo (1). Em todas as especificações foram utilizadas dummies de tempo. Além disso, nas estimações feitas por MGM, as variáveis *desig*, *lnpibp*, *estudo*, *urb*, *crescpib* e *lnseg* foram tratadas como potencialmente endógenas. Cada coeficiente estimado pode ser interpretado como a mudança relativa do crime causado por um aumento unitário na variável explicativa correspondente. Dado que o modelo inclui uma variável dependente defasada, cada coeficiente estimado corresponde ao efeito de curto prazo de sua respectiva variável. O efeito de longo prazo pode ser obtido dividindo o efeito de curto prazo por 1 menos o coeficiente da variável dependente defasada.

⁸Assume-se que o erro v_{it} é homocedástico e não auto-correlacionado.

⁹Ver Blundell e Bond (1998).

Tabela 2 - Resultados das estimações

	[a]	[b]	[c]	[d]	[e]	[f]	[g]
	MQO	MQO	WG	MGM-DIF	MGM-SIS	MGM-SYS	MGM-SIS
$\ln\text{razhom}_{i,t-1}$	-	0,5687*	0,0635**	0,0418*	0,4893*	0,4548*	0,4094*
	(0,0035)	(0,0063)	(0,0272)	(0,0134)	(0,0910)	(0,0915)	(0,1093)
desig	0,0121*	0,0063**	0,0039**	0,0016	0,0067**	0,0097*	0,0092**
	(0,0035)	(0,0026)	(0,0015)	(0,0016)	(0,0031)	(0,0033)	(0,0040)
$\ln\text{pibp}$	0,7586*	0,3601*	0,8301*	0,9609*	0,4132*	0,4478*	0,4469*
	(0,0298)	(0,0854)	(0,0283)	(0,0382)	(0,0730)	(0,0574)	(0,0726)
estudo	-0,1031*	-0,0449**	-0,0320**	0,0095	-0,0639*	-0,0544**	-0,0808**
	(0,0198)	(0,0197)	(0,0131)	(0,0098)	(0,0195)	(0,0235)	(0,0303)
urb	-0,0269*	-0,0123*	0,0030	0,0081	-0,0142*	-0,0174*	-0,0169*
	(0,0016)	(0,0025)	(0,0034)	(0,0137)	(0,0032)	(0,0033)	(0,0038)
crescpib	0,019	0,0008	-0,0034*	-0,0044*	-0,0008	-0,0025***	-0,0027***
	(0,0021)	(0,0013)	(0,0007)	(0,0005)	(0,0009)	(0,0012)	(0,0014)
Inseg	0,0255	0,0100	0,0231***	0,0068	0,0006	-0,003	0,0155
	(0,0259)	(0,0168)	(0,0129)	(0,0155)	(0,0256)	(0,0271)	(0,0329)
constante	14,4029*	5,8865*	10,2211*	-	7,2881*	7,8201*	8,6256*
	(0,2792)	(1,4451)	(0,4679)	-	(1,4914)	(1,6766)	(1,8432)
teste de Sargan (p-valor)	-	-	-	1	1	1	1
correlação serial (p-valor)							
Primeira ordem	0,00	0,00	0,000	0,02	0,07	0,08	0,06
Segunda ordem	0,00	0,00	0,000	0,22	0,40	0,38	0,40
Correção para erro de medida	não	não	não	não	não	sim	sim
Defasagens dos instrumentos	-	-	-	2 a infinito	2 a infinito	3 a infinito	3 a 4

Fonte: DATA-SUS. Elaboração própria.

Nota: A variável $\ln\text{razhom}$ é usada como variável dependente. Em parênteses são apresentados os respectivos desvios-padrão. Em todas as especificações foram utilizadas dummies de tempo. Os sinais *, **, *** representam, respectivamente, significância de 1%, 5% e 10%.

Na coluna [a], o modelo foi estimado por MQO sem a variável dependente defasada e controles para o efeito fixo de cada país ao longo do tempo. Com exceção dos coeficientes da taxa de urbanização (urb) e dos gastos em segurança pública per capita (seg) que revelaram ser respectivamente negativo e insignificante, as demais variáveis apresentaram sinais conforme a teoria. Contudo, o modelo é rejeitado pelos testes de correlação serial de primeira e segunda ordem indicando, possivelmente, omissão de variáveis cujo impacto é persistente ao longo do tempo ou falta de controle para o efeito dinâmico da criminalidade.

Adicionando a variável dependente defasada como variável explicativa (coluna [b]), observamos que seu estimador MQO apresentou um coeficiente positivo e significância ao nível de 1%. Esse resultado comprova que a criminalidade tem um componente inercial. O sinal e a significância das demais variáveis no modelo não se alteraram. Contudo, os testes de especificação continuam indicando a presença de correlação serial. Desta forma, existe a necessidade de se controlar o efeito fixo de cada estado ao longo do tempo.

Na coluna [c], o modelo foi estimado por WG como forma de se controlar o efeito fixo de cada estado ao longo do tempo. O coeficiente da variável dependente defasada apresentou um valor significativo relativamente mais baixo do que o estimado por MQO. A taxa de urbanização, nesta especificação, deixou de ser significativa. Além disso, os estimadores para crescpib e seg passaram a ser significantes com sinais respectivamente negativo e positivo.

Conforme discutido na seção anterior, WG forneceria um estimador negativamente viesado, para amostras com períodos de tempo pequeno, enquanto o coeficiente estimado por MQO seria positivamente viesado. Além disso, estes métodos não levam em conta problemas de endogeneidade das variáveis explicativas devido a causalidade inversa. Uma forma de se evitar esse problema é a estimação por MGM-DIF.

Somente os estimadores das variáveis *desig* e *crescpib* apresentaram significância com sinais respectivamente positivo e negativo na estimação por MGM-DIF (coluna [d]). O coeficiente da variável dependente defasada ficou bem abaixo do estimador por WG sugerindo problemas nos instrumentos. Note, contudo, que esta estimação se mostrou robusta aos testes de especificação.

Na coluna [e] são apresentados os resultados da estimação por MGM-SIS. O coeficiente estimado da variável dependente defasada foi de 0,48, um número menor do que o fornecido por MQO e maior que obtido por WG. Nem o teste de Sargan e nem os testes de correlação serial detectaram problemas nos instrumentos. Desta forma, o MGM-SIS permite um ganho considerável de precisão comparado ao MGM-DIF. O resultado dessa regressão comprova que os fatores econômicos que determinam a criminalidade são o nível de criminalidade do período anterior, o grau de desigualdade de renda, o PIB per capita, o nível de escolaridade e o grau de urbanização. Os gastos em segurança pública per capita e o crescimento do PIB não apresentaram significância.

A dinâmica da criminalidade pode ser entendida pela queda do custo de se cometer o crime novamente. A atividade ilegal, como em qualquer outra profissão, gera um ganho de habilidade ao longo da jornada de trabalho. Dessa forma, o custo de se cometer um crime e o valor moral do criminoso diminuí com o tempo de "trabalho". Além disso, altas taxas de criminalidade no período anterior favorece a crença de um sistema judicial frágil e impotente estimulando assim a prática da atividade criminosa no período atual.

O estimador para a desigualdade de renda apresentou um coeficiente significativo de 0,006. Esse resultado indica que indivíduos, em geral, por não terem uma expectativa de aumento do padrão de vida através de atividades legais, têm incentivos a cometerem crimes. Contudo um estudo mais detalhado sobre os impactos da desigualdade de renda sobre a criminalidade deve ser feito tendo em vista os diversos índices de desigualdade de renda propostos pela literatura.

A variável gastos em segurança pública per capita (*seg*) não apresentou significância. A má utilização dos recursos públicos e a falta de um órgão nacional coordenando as atividades das secretarias de segurança estaduais seja a melhor explicação para este resultado.

Um ano a mais de estudo mostrou que pode procovar uma queda de 6% na taxa de criminalidade no curto prazo e de, aproximadamente, 12% no longo prazo. Dessa forma, os benefícios gerados pela educação como aumento do valor moral do indivíduo e melhores oportunidades de emprego são maiores do que a queda do custo de se cometer um crime e da probabilidade de ser preso.

O PIB per capita apresentou uma elasticidade significativa de 0,41 sobre a criminalidade. Embora uma renda per capita mais alta represente maiores oportunidades de emprego, o ganho obtido com na atividade ilegal é ainda mais elvada.

Por fim, a taxa de urbanização apresentou um coeficiente significativo negativo. Segundo Glaeser e Sacerdote (1999), a facilidade na fuga e a difícil identificação do criminoso incentiva a criminalidade em áreas mais densamente povoadas. Contudo, obteve-se um efeito contrário.

Como forma de controlar para erros de medida, cada uma das variáveis foi tratada, como discutido na seção anterior, em função também de um choque aleatório que varia de acordo com estado e tempo e um efeito fixo. Desta forma, os instrumentos em $t - 2$ para as equações em primeira diferença e em $t - 1$ para as equações em níveis são inválidos. Na coluna [f], o modelo é estimado por MGM-SIS levando-se em conta este problema. De modo geral, com exceção da variável *crescpib*, a significância e o sinal dos estimadores pouco se alterou com a exclusão dos instrumentos citados. O teste de Sargan e os de correlação serial indicam que os instrumentos são válidos. A taxa de criminalidade se mostrou contra-cíclica em relação ao crescimento do PIB apesar do PIB per capita continue sendo negativamente significativo.

Embora o uso do conjunto total de instrumentos possíveis aumente a eficiência do estimador, para amostras finitas, o estimador para α_1 torna-se viesado¹⁰. Como forma de evitar este tipo de problema, na coluna [g] o modelo foi estimado considerando apenas os instrumentos mais recentes com o número máximo de duas defasagens. Os resultados obtidos são similares aos da coluna [f].

6 Conclusões

A partir de dados de painel para estados brasileiros no período 1984-1998, estimamos os determinantes da taxa de criminalidade brasileira utilizando o método generalizado dos momentos em sistema proposto por Blundell e Bond (1998). Desta forma, foi possível amenizar problemas econométricos, que afetam a maioria dos trabalhos nesta área, como endogeneidade das variáveis explicativas e erros de medida gerados pela subnotificação das taxas de crimes. As estimativas obtidas permitem concluir que o grau de desigualdade de renda e a taxa de criminalidade do período anterior geram um efeito positivo sobre a taxa de criminalidade do período presente, enquanto que o PIB per-capita, o nível de escolaridade, o grau de urbanização e o crescimento do PIB têm efeitos negativos. Além disso, o uso do método generalizado dos momentos em primeira diferença (Arellano e Bond, 1991), nesse caso, não revelou ser apropriado, devido a presença de instrumentos fracos, apresentando um estimador para a taxa de criminalidade defasada bem abaixo do estimador *Within-Groups*. O uso de instrumentos adicionais em primeira diferença para as equações em níveis através do MGM-SIS mostrou ser mais informativo. Em todas as estimações por MGM, os testes de Sargan e de correlação serial mostraram que os instrumentos são válidos.

¹⁰Ver Judson e Owen (1996).

Referências

- [1] Araújo, Ari Franco de e Fajnzylber, Pablo (2001). "O que causa a criminalidade violenta no Brasil? Uma análise a partir do modelo econômico do crime: 1981 a 1996". mimeo.
- [2] Arellano, Manuel e Bond, Stephen (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations". *Review of Economic Studies*, 58, 277-297.
- [3] Arellano, Manuel e Bover, Oliver (1995). "Another look at the instrumental variable estimation of error-components models". *Journal of Econometrics*, 68, 29-52.
- [4] Becker, Gary S. (1968). "Crime and punishment: an economic approach". *Journal of Political Economy*, 76, 169-217.
- [5] Blundell, Richard e Bond, Stephen (1998). "Initial conditional and moment restrictions in dynamic panel data models". *Journal of Econometrics*, 87(1), 115-143.
- [6] Blundell, Richard e Bond, Stephen (2000). "GMM estimation with persistent panel data: an application to production functions". *Econometric Reviews*, 19(3), 321-340.
- [7] Blundell, Richard, Bond, Stephen e Windmeijer, F. (2000). "Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator". In B. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Elsevier Science.
- [8] Bond, Stephen, Hoeffler, Anke e Temple, Jonathan (2001). "GMM estimation of empirical growth models". mimeo.
- [9] Hsiao, C. (1986). *Analysis of panel data*. Cambridge: Cambridge University Press.
- [10] Fajnzylber, Pablo, Lederman, Daniel e Loayza, Norman (2000). "Crime and victimization: an economic perspective". Mimeo.
- [11] Fajnzylber, Pablo, Lederman, Daniel e Loayza, Norman (2000). "What causes violent crime". *European Economic Review*, 46, 1323-1357.
- [12] Fajnzylber, Pablo, Lederman, Daniel e Loayza, Norman (2002). "Inequality and violent crime". *Journal of Law and Economics*.
- [13] Fleisher, B. M. (1966). "The effect of income on delinquency". *American Economic Review*, 56, 118-137.
- [14] Cerqueira, Daniel e Lobão, Waldir (2003). "Condições socioeconômicas, polícia e produção criminal". IPEA, 2003, mimeo.

- [15] Glaeser, E. e Sacerdote, B. (1999). "Why is there more crime in cities?". *Journal Political Economy*, 107, 225-258.
- [16] Glaeser, E., Sacerdote, B. e Scheinkman, J. (1996). "Crime and social interations". *Quarterly Journal of Economics*, 111, 507-548.
- [17] Ehrlich, Isaac (1973). "Participation in illegetimate activities: a theoretical and empirical inves-tigation". *Journal of Political Economy*, 81, 521-565.
- [18] Ehrlich, Isaac (1975). "On the relation between education and crime". *Education, income and human behavior*. McGraw-Hill: New York.
- [19] Judson, Ruth e Owen, Ann (1996). "Estimating dynamic panel data models: a practical guide for macroeconomics". Federal Reserve Board of Governors.
- [20] Levitt, Steven (1996). "The effect of prisison population size on crime rates: evidence from prisison overcrowding litigation". *Quarterly Journal of Economics*, 111, 319-352.
- [21] Levitt, Steven (1997). "Using electoral cycles in police hiring to estimate the effect of police on crime". *American Economic Review*, 87, 270-290.
- [22] Levitt, Steven (1998). "Why do increased arrest rates appear to reduce crime: deterrence, inca-pacitation or measurement of error?". *Economic Inquiry*, 36, 353-372.
- [23] Nickell, Stephen (1981). "Biases in dynamic models with fixed effects". *Econometrica*, 79(6): 1417-1426.