

ESTUDO DA CRIMINALIDADE VIOLENTA NA CIDADE DO RECIFE: O ESPAÇO REALMENTE É RELEVANTE?

Giuseppe Trevisan Cruz (PIMES/UFPE)
Tatiane Almeida de Menezes (PIMES/UFPE)

RESUMO

Um segmento importante da literatura de Economia do Crime afirma que, além das variáveis socioeconômicas, o espaço é fator fundamental a se associar com a criminalidade. Dada a falta de evidências empíricas sobre a relação entre espaço e crime, este estudo tem por objetivo identificar as correlações entre as variáveis ambientais com a taxa de homicídio nos bairros da cidade do Recife. Para isso, foram construídas variáveis de ambiente que capturam características relacionadas à distribuição dos tipos de domicílios dos bairros do Recife e foi implementada a técnica de econometria espacial para averiguar efeitos de *spillover* espaciais. O modelo SAR apresenta o melhor ajuste e mostra que a proporção de estabelecimentos não-residenciais em relação ao total de estabelecimentos do bairro tem uma relação positiva com a criminalidade e a concentração de domicílios residenciais está associada a índices de criminalidades mais baixos. As correlações das variáveis socioeconômicas seguem o padrão da maioria dos achados da literatura nacional e internacional, exceto para o caso da densidade demográfica.

Palavras-Chave: Economia do Crime, Criminalidade, Econometria Espacial, Testes de Especificação Econométrica, Spatial Autoregressive Model (SAR).

ABSTRACT

A stream of studies of Criminology says that space is a fundamental factor to associate with crime. Given the lack of empirical evidence on the relationship between space and crime, this study aimed to identify the correlations between environmental variables with the homicide rate in Recife's neighborhoods. We built environment variables that capture the characteristics related to the distribution of the types of households of the Recife's neighborhoods and implemented spatial econometric's techniques to determine spatial spillover effects. The SAR model has the best fit and shows that the proportion of non-residential establishments in relation to the total number of establishments in the neighborhoods has a positive relationship with neighborhood crime and the concentration of residential homes is associated with lower rates of criminalities. Correlations of socioeconomic variables follow the pattern of most national and international literature findings, except for the demographic density.

Keywords: Economics of Crime, criminality, Spatial Econometrics, Econometric Specification tests, Spatial Autoregressive Model (SAR).

INTRODUÇÃO

A criminalidade é um fenômeno que tem se tornado fonte de estudo de várias ciências sociais. No campo da economia do crime, Becker (1968) foi o primeiro a modelar o comportamento do agente com base na relação custo/benefício em se praticar o ato ilícito. A

partir desse estudo de referência, diversos trabalhos procuraram capturar como os fatores socioeconômicos se relacionam com o crime (MESSNER, 1982; MESSNER e ROSENFELD, 1997; LAFREE, 1999; WITTE 1980; GLAESER e SACERDOTE, 1999; FAJNZYLBER et al., 2002; PRIDEMORE, 2011; SOARES, 2000). Os achados empíricos são muitas vezes controversos e não há um consenso na literatura sobre os sentidos de parte dessas relações.

Há também uma vertente de pesquisas em criminologia que aborda a importância do espaço na determinação da atividade criminal e como as características do ambiente podem influenciar a atmosfera em que o crime ocorre. Brantingham e Brantingham (1995) atentam para o fato de que existe uma série de questões que tornam importante compreender como o ambiente e seus elementos contribuem para a escolha das vítimas e dos locais destinados ao crime por parte dos delinquentes, questões tais exploradas por uma gama de autores (BRANTINGHAM e BRANTINGHAM, 1993b; WILCOX, 1973; BRANTINGHAM et al., 1977; BRANTINGHAM et al., 1995; CLARKE e MAYHEW, 1994; POYNER, 1992; ECK e SPELLMAN, 1994; FLEMING et al., 1994). Segundo Brantingham e Brantingham (1996), há locais que criam oportunidades à ocorrência de algum tipo de crime de acordo com suas características físicas peculiares ou com o perfil dos potenciais alvos.

A lacuna de informações a respeito de uma influência mais explícita do ambiente na criminalidade é muito grande. Dada essa motivação, este artigo procura responder se realmente o ambiente é fator relevante para explicar a criminalidade e de que forma ambos se relacionam. Os indícios de que o espaço é relevante para explicar o crime são muitos, tendo em vista que, através de comprovações empíricas, sabe-se que a distribuição espacial e temporal do crime não é aleatória, apresentando assim um padrão conforme o ambiente (WORTLEY e MAZEROLLE, 2008). A dificuldade em realizar trabalhos deste tipo está na ausência de banco de dados que incorporem informações mais específicas sobre o espaço, e um dos diferenciais desta pesquisa está na utilização de uma plataforma de dados que torna possível trazer tais características ambientais no modelo econométrico. Outra peculiaridade deste trabalho é o tratamento de dados em elevado nível de desagregação (bairros), o que permite capturar os efeitos de transbordamento entre vizinhos com maior precisão.

O objetivo deste trabalho é capturar as correlações entre variáveis ambientais com a taxa de homicídio na cidade do Recife, já que carecemos de uma estratégia de identificação forte que venha corrigir a endogeneidade existente por variáveis omitidas. Com a inserção das características ambientais por bairro no modelo econométrico, o estudo se torna inovador e busca contribuir com a compreensão do papel do espaço na criminalidade utilizando a metodologia de econometria espacial, a fim captar possíveis influências de características de bairros vizinhos nas taxas de homicídios. Obtivemos, através da seleção de um modelo – com embasamento estatístico - que descreve a relação entre crime e as variáveis explicativas, os efeitos diretos, indiretos e totais, mesmo cientes de que o viés persiste¹. As variáveis de ambiente foram construídas com base nas informações contidas no banco de dados do CNEFE (Cadastro Nacional de Endereços para Fins Estatísticos). De posse de uma base de dados única fornecida pela SDS/PE (Secretaria de Segurança Pública do estado de Pernambuco) e pelo Censo Demográfico do IBGE (Instituto Brasileiro de Geografia e Estatística) foi possível construir as taxas de homicídio e obter as variáveis socioeconômicas, todas em nível de bairro.

No tocante ao Brasil, existem trabalhos que utilizam a técnica de econometria espacial para analisar o crime com os fatores socioeconômicos. Menezes et al. (2011) desenvolveram um estudo para os bairros da cidade do Recife que procura identificar os efeitos da pobreza e da desigualdade nas taxas de homicídios, além de procurar entender a relação da criminalidade entre vizinhos. Almeida et al. (2005) também contribuem com análise espacial

¹ Na seção de Análise de Resultados discutimos em detalhes a questão da endogeneidade no modelo, procurando mostrar que não há problemas de causalidade reversa, apenas viés de variáveis não-espaciais omitidas.

investigando a criminalidade nos municípios de Minas Gerais e detectam a importância do espaço para explicar as ocorrências de crime no estado. Da mesma forma, Scorzafé e Soares (2009) utilizaram econometria espacial em seu estudo aplicado para a cidade de São Paulo, mas verificaram a relação entre desigualdade de renda e criminalidade, encontrando um efeito positivo sobre a taxa de homicídio.

O trabalho está dividido em seis seções, incluindo esta introdução. A próxima seção contém informações sobre os bancos de dados utilizados e a análise descritiva dos mesmos. Em seguida, uma breve contextualização em relação à criminalidade da cidade do Recife é feita, motivando o estudo com algumas informações sobre os homicídios e o ambiente espacial da cidade. A quarta seção apresenta a metodologia utilizada para implementar a estratégia empírica, contendo toda a estrutura econométrica do trabalho. Por fim, a seção de Análise de Resultados apresenta as regressões realizadas e a discussão do trabalho, seguido da conclusão.

DADOS E ANÁLISE DESCRITIVA

Para realizar as estimações foram utilizados três bancos de dados distintos. As informações referentes a crimes violentos letais e intencionais foram obtidas através do banco de dados INFOPOL SDS/PE, da Secretaria de Defesa Social de Pernambuco, e constam de um histórico de homicídios (e latrocínios) no período de 2004 a 2011 na cidade do Recife, registrados por bairro de ocorrência. Os crimes são descritos por dia da semana, dia do mês e ano e turno do dia em que ocorreu o homicídio. Além do período e do local geográfico que referencia onde ocorreu o crime, o banco traz um conteúdo completo sobre a vítima com características de raça, idade, sexo e também sobre o delito, como objeto utilizado pelo agressor e motivação do crime. Para este estudo, foi realizado um corte transversal no período de 2010 devido à falta de informações socioeconômicas e ambientais em outros anos, tendo como unidade de observação os bairros do Recife.

As variáveis socioeconômicas utilizadas foram coletadas do Censo Demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE) para o período de 2010 e por bairro, a saber: população total, renda *per capita* e área do bairro. A partir dessas variáveis, foram construídas outras para analisar a criminalidade e que foram utilizadas em outros trabalhos na literatura de economia do crime, como: índice de Gini, população e crescimento populacional. O Gini é o índice de desigualdade mais utilizado na literatura de economia do crime e, de acordo com Hoffmann (2001), detém a propriedade de ser mais sensível a transferências regressivas no intervalo onde a densidade de frequência é maior, geralmente em torno da moda ou mediana da distribuição. O índice foi calculado por bairro, considerando a renda *per capita* dos indivíduos residentes num determinado bairro.

A definição da amostra do Censo de 2010 pelo IBGE se dá com base em um banco de dados denominado Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) que registra e caracteriza todos os endereços do país. Com ele, é possível inserir informações ambientais que irão corroborar para captar algum tipo de correlação entre os homicídios e as características dos bairros. As observações são por domicílio e não são considerados lotes vagos ou edificações em construção sem morador. Os componentes do endereço são identificados por logradouro, número da edificação, complemento, localidade (bairro) e ponto de referência. Pode-se quantificar e identificar no banco os tipos de estabelecimento por bairro, como por exemplo, saber se um endereço está associado a um local destinado à moradia ou a alguma atividade comercial.

Com posse dos dois bancos de dados do IBGE foram construídas quatro variáveis ambientais. Como detalhado na Tabela 1, a variável “naores” diz respeito à proporção de estabelecimentos não residenciais no bairro “i” em relação ao total de estabelecimentos desse

mesmo bairro. As outras variáveis correspondem à razão da quantidade de estabelecimentos não-residenciais (prop.nres), residenciais (prop.res) e totais (prop.tot) com a área do bairro medida em quilômetros quadrados. Elas foram concebidas dessa maneira para transmitir uma ideia de como os estabelecimentos estão distribuídos no espaço - ou seja, uma densidade desses estabelecimentos -, traçando, assim, um perfil do local.

Ainda na Tabela 1 são apresentadas as estatísticas descritivas das variáveis que foram utilizadas no modelo empírico, cada uma com 94 observações que correspondem a cada bairro da cidade. A primeira delas é a taxa de homicídio² (tx_h) por bairro do Recife, sendo a quantidade de mortes dividida pela população multiplicada por 100.000. A média da taxa de homicídios é alta (78,642) porque existem duas observações aberrantes no banco, que são os bairros do Recife Antigo e Santo Antônio³, ambos localizados no centro da cidade, com taxas anuais de homicídio de 332,22 e 3.157,9 respectivamente - a taxa média de homicídio sem esses dois bairros cai para 42,415.

Tabela 1 – Análise Descritiva das Variáveis Socioeconômicas e Ambientais

Variáveis		Média	Desvio Padrão	Mínimo	Máximo
tx_h	(taxa de homicídio)	78,642	324,797	0	3.157,9
dIndens	(tx de cresc. da dens. demo.)	0,16	0,42	-0,172	2,87
dens2010 (km ²)	(densidade demográfica)	12.566,59	7.068,259	139	31.970
gini	(índice de Gini)	0,481	0,076	0,334	0,787
renda	(renda <i>per capita</i>)	279,00	296,74	47,19	1.333,89
naores	(prop. de estab. não-residenciais sobre o total de estab.)	0,137	0,116	0	83,49
prop.nres	(densidade de estab. não-residenciais)	521,67	301,87	0	1.623,51
prop.res	(densidade de estab. residenciais)	4.058,31	2.345,13	11,251	10.514,83
prop.tot	(densidade total de	4.579,98	2.543,26	11,251	11.986,04

² Na construção da taxa de homicídio foram considerados os latrocínios e homicídios, que se refere aos crimes violentos letais e intencionais. Apenas com o propósito de facilitar a discussão, foi adotado o termo “taxa de homicídio”.

³ São bairros pobres, com baixo número de domicílios residenciais e com atividade noturna intensa, de modo que poucas ocorrências dispare a taxa de homicídio.

Total de observações: 94

Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE, Censo Demográfico e CNEFE (IBGE).

A densidade demográfica (dens2010) da cidade é moderada, possuindo cerca de 12.566 habitantes por km², em média. Além do fato de haver pequenos bairros onde as comunidades tem uma concentração de pessoas muito grande (com adensamento horizontal), essa alta densidade também decorre da verticalização da cidade e do mercado imobiliário bastante aquecido - já que um número muito grande de pessoas se instala na cidade devido ao bom crescimento da atividade econômica em vários setores. Também é interessante observar que o inchaço da cidade parece ser maior por causa de vazios urbanos decorrentes de condições geográficas, como áreas de mangue e verde. Pau Ferro é o bairro que apresenta menor densidade, seguido do Recife Antigo e Guabiraba, onde o último possui a maior área territorial, situado no extremo norte da cidade. Brasília Teimosa, Alto José do Pinho e Mangueira apresentam as maiores densidades demográficas (adensamento horizontal) e estão entre os 30 menores bairros em termos de área geográfica. A taxa de crescimento populacional dos bairros (dlndenss) no período entre 2000 e 2010 foi, em média, de 16%.

Em termos de desigualdade de renda (gini), os bairros do Recife apresentam uma dispersão moderada em torno da média (0,48), já que 0,07 pontos em termos de índice de Gini é algo significativo. Cabe ressaltar que se encontram muitas comunidades pobres vizinhas a áreas com padrões financeiros mais elevados dentro de um mesmo bairro, caracterizando também uma heterogeneidade intrabairro. Pau Ferro (0,726) e Santo Antônio (0,654) se enquadram nos bairros mais desiguais da cidade, onde o primeiro é formado de uma população de granjeiros e comerciantes – sendo considerado um reduto rural - e, o segundo, fica no centro da cidade e contém uma série de pontos históricos e áreas não residenciais. Os bairros de Jaqueira e Graças são os menos desiguais, com gini de 0,334 e 0,337, estando entre os quatro mais ricos da cidade em termos de renda *per capita* – R\$1.333,00 e R\$1.066,00 respectivamente. Além dos dois últimos mencionados, encontramos os bairros de Casa Forte (R\$1.111,00 *per capita*) e Afogados (R\$1.077,00) com elevada renda *per capita*, mas é curioso como são heterogêneos entre si. Afogados é o bairro mais populoso e o que detém a maior parcela de estabelecimentos não-residenciais dentre os quatro, onde a elevada renda *per capita* pode ser atribuída ao alto nível de atividades ligadas ao comércio e às indústrias na região. O bairro das Graças possui um polo de saúde desenvolvido que movimenta muitos recursos, enquanto Jaqueira e Casa Forte são bairros mais destinados a residências, embora tenha um setor de serviços bem aquecido. A média de R\$279,00 de renda *per capita* é bastante baixa para a cidade e tem uma dispersão alta (R\$296,00) devido a alguns pontos atípicos. Além disso, cerca de 40% dos bairros da cidade possuem renda *per capita* abaixo de R\$100,00.

A cidade do Recife é composta por uma porção de escolas e faculdades, hospitais e clínicas de saúde, estabelecimentos religiosos e comerciais em geral, dentre vários outros tipos de locais não destinados à moradia, e esses representam em média 13% (naores) dos estabelecimentos dos bairros, com desvio padrão de 11%. Santo Antônio é o bairro que apresenta a maior proporção de estabelecimentos não-residenciais (83%), seguido do Recife Antigo (67%) e Boa Vista (21%), sendo observações bastante atípicas em relação à distribuição dos bairros - os três estão localizados no centro da cidade. Pau Ferro, Aflitos, Monteiro e Boa Viagem são os mais residenciais e situados em pontos bem diferentes no mapa da cidade.

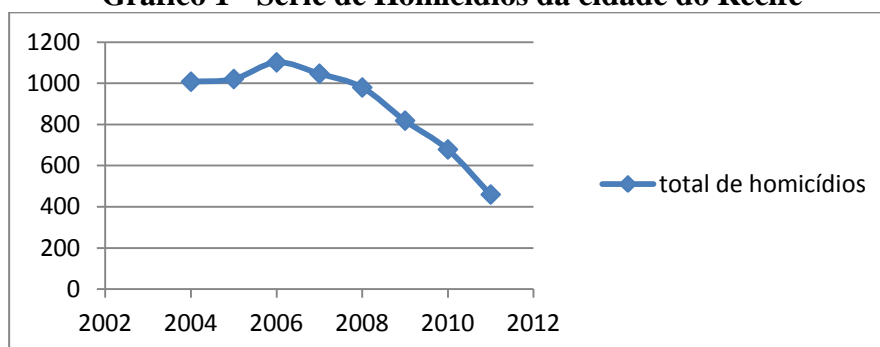
O cenário assume um aspecto diferente se se considera a concentração por tipo de estabelecimento nos bairros. Os 4 bairros que apresentam a quantidade de domicílios de

forma menos concentrada (prop.res) na cidade são, respectivamente: Pau Ferro, Guabiraba, Cidade Universitária, Dois Irmãos. Analisando descritivamente os bairros que possuem a maior concentração de estabelecimentos não-residenciais (prop.nres), tem-se: Santo Antônio (1.623,51), Brasília Teimosa, Boa Vista e Totó. Os bairros de Brasília Teimosa (11.986,04 domicílios), Alto José do Pinho e Mangueira detêm a maior concentração de domicílios (prop.tot), embora Boa Viagem, Várzea e Cohab possuam a maior quantidade de endereços.

A CRIMINALIDADE NA CIDADE DO RECIFE: O ESPAÇO É IMPORTANTE?

A cidade do Recife, capital do estado de Pernambuco e situada na região Nordeste do Brasil, é composta por 94 bairros bastante heterogêneos, tanto em questões sociais quanto econômicas, e considerada uma capital com alto nível de criminalidade. Mesmo sendo uma cidade com alto nível de periculosidade, a série representada pelo Gráfico 1 mostra que os homicídios (em termos absolutos) vêm decaindo ao longo do tempo, desde 2005 - em termos de taxa de homicídio, essa redução também é significativa tendo em vista que a cidade obteve um crescimento populacional de 2% (de 2000 a 2010). De 2007 até 2012, houve uma redução de 34% na taxa de crimes violentos letais intencionais de acordo com os dados da Secretaria de Defesa Social.

Gráfico 1 - Série de Homicídios da cidade do Recife



Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE

No ano de 2010, houve 678 homicídios na cidade, conforme mostra a Tabela 2. A maior proporção de homicídios foram cometidos com arma de fogo (cerca de 86%), o que denota a facilidade de acesso a algum tipo de arma de disparo na cidade. Do total, 63% dos homicídios aconteceram no período da noite e madrugada, momento em que as ruas estão menos movimentadas em comparação com o período diurno onde o fluxo de pessoas é bem maior devido às suas atividades rotineiras - trabalho, escola, faculdade, entre outras. O movimento de pessoas no período noturno dá-se, em sua maior parte, por motivos de entretenimento, compreendendo ambientes como bares/restaurantes, cinemas e casas noturnas. E, ainda de acordo com a Tabela 2, observa-se que aproximadamente 68% dos homicídios em 2010 ocorreram na sexta, no sábado, no domingo e na segunda⁴, dias em que os indivíduos estão mais inclinados às atividades de lazer. Cerca de 74% dos homicídios acontecem em logradouro público.

Tabela 2 - Dados de homicídios da cidade do Recife em 2010

TOTAL	FREQÜÊNC IA	TOTAL	FREQÜÊNC IA
-------	----------------	-------	----------------

⁴ Muitas pessoas emendam a noite do domingo com a madrugada da segunda-feira em suas atividades de lazer, movimentando bares e restaurantes. É nesse período em que ocorre a maior parte dos homicídios do dia.

Objeto utilizado na Cena do crime			Dia da Semana de ocorrência		
Arma branca	46	6,78%	Segunda	111	16,37%
Arma de fogo	584	86,14%	Terça	84	12,39%
outras	48	7,08%	Quarta	69	10,18%
Período do dia			Quinta	58	8,55%
Madrugada	176	25,96%	Sexta	97	14,31%
Manhã	92	13,57%	Sábado	124	18,29%
Tarde	159	23,45%	Domingo	135	19,91%
Noite	251	37,02%			

Total de Homicídios: 678

Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE

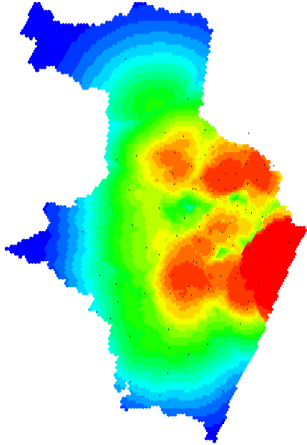
Importante notar que o Recife também apresenta heterogeneidade em relação à forma de ocupação do espaço geográfico dos bairros, caracterizando por apresentar zonas com muita concentração de domicílios e também por algumas áreas de vazio urbano. No total, são 581.136 endereços registrados na cidade no ano de 2010, onde 88% deles são destinados à moradia, devendo-se atentar ao fato de que se podem ter vários domicílios em um mesmo loteamento (um prédio, por exemplo).

Se analisarmos as distribuições dos homicídios e dos estabelecimentos que são destinados a qualquer fim diferente de moradia particular fica notória a confluência do adensamento de ambas. As Figuras 1 e 2, que constam de um mapa Kernel, mostram que os clusters de homicídios da cidade do Recife estão localizados, praticamente, nas mesmas áreas onde se encontram maior concentração de estabelecimentos não-residenciais. As Figuras 3 e 4 mostram a distribuição da densidade demográfica e da densidade de estabelecimentos residenciais (razão entre o total de residências e a área do bairro), e pode-se perceber que estão muito correlacionadas (correlação de 92%). Traçando uma analogia entre as figuras, nota-se que o padrão espacial da Figura 1 é mais semelhante com o da Figura 2 do que com os padrões das Figuras 3 e 4. Menezes et al (2011) sugerem que um dos padrões espaciais associados à criminalidade é devido às áreas de subúrbio, que apresentam alta densidade e as vítimas são feitas perto de suas residências. Além disso, as figuras denotam que existe um fator ambiental que pode estar correlacionado com a criminalidade e que antes não era observado.

Esses indícios captados pelas distribuições Kernel motivam a incorporação do ambiente no estudo da criminalidade para os bairros do Recife. As Figuras 2 a 4 sugerem que a ocorrência de crime não se dá de forma aleatória no espaço. Além de questões socioeconômicas, o ambiente também se configura como fator responsável por certos padrões de criminalidade em determinadas áreas, conforme Brantingham e Brantingham (1995). Pode ser que a característica ambiental e até socioeconômica de um determinado bairro influencie, de alguma forma, a criminalidade de um bairro vizinho, visto que os *spillovers* são mais facilmente captados quando se trata de um nível de desagregação tão baixo. A ausência de vigilância em função de poucos estabelecimentos residenciais gera o ambiente propício para a existência de atos ilícitos. Em função disto, a literatura recente de econometria espacial mensura o impacto dos efeitos diretos e indiretos dos fenômenos socioambientais (LESAGE e PACE, 2009). É provável que a proporção de estabelecimentos não-residenciais tenha um efeito significativo na criminalidade do próprio bairro (efeito direto) e, de alguma forma, afete a ocorrência de homicídios em outros bairros vizinhos (efeito indireto), por exemplo. A seção

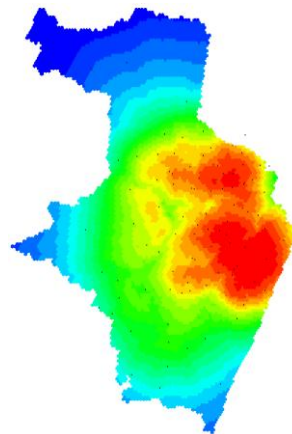
seguinte apresenta a metodologia utilizada para averiguar o tratamento das variáveis ambientais e a dependência espacial entre as unidades de observação e detectar sua influência na criminalidade no Recife.

Figura 1 - Taxas de homicídio no Recife em 2010



Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE

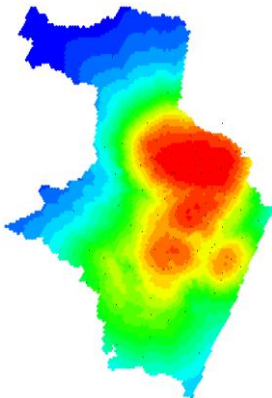
Figura 2 - Proporção de estabelecimentos não-residenciais no Recife em 2010



Fonte: Elaborado pelo

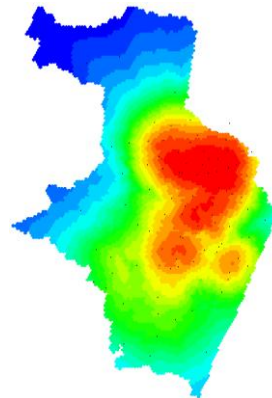
CNEFE/IBGE

Figura 3 - Densidade demográfica no Recife em 2010



Fonte: Elaborado pelo Autor a partir das informações do Censo Demográfico (IBGE)

Figura 4 - Densidade de domicílios⁵ no Recife em 2010



Fonte: Elaborado pelo

CNEFE/IBGE

METODOLOGIA

A literatura de econometria espacial busca implantar nos modelos de regressão linear a influência das características das observações vizinhas numa localidade próxima. Essa dependência em termos de espaço pode ser verificada na variável dependente, na(s) variável(eis) independente(s), na disturbância ou até mesmo em mais de um tipo de variável simultaneamente. A incorporação da dependência espacial no modelo é feita por meio de uma matriz de pesos denominada **W**, onde tais pesos são definidos de acordo com a importância que se atribui à uma observação vizinha. O mais comum na literatura é a construção de pesos espaciais a partir das características geográficas das unidades observadas, e tais pesos se

⁵ Quantidade de domicílios destinados à moradia dividida pela área do bairro.

reduzem à medida que a unidade se torna cada vez mais distante de uma referência (TYSZLER, 2006).

Após a escolha da matriz de pesos vem a configuração da forma funcional. No modelo clássico, tem-se que uma determinada característica está correlacionada com um conjunto de regressores mais um termo aleatório, ou seja:

$$y = \mathbf{X}\beta + u, \quad (1.1)$$

onde “y” é o vetor da variável dependente, \mathbf{X} é a matriz das variáveis independentes e “u” o termo de erro. Se houver algum tipo de dependência espacial, tanto nos regressores quanto no regressando, os estimadores ($\hat{\beta}$) de MQO serão viesados e o problema de endogeneidade emerge – viés de variável espacial omitida. Caso a influência espacial esteja concentrada apenas no erro, os estimadores de MQO não serão viesados, mas também não serão consistentes (ERTUR e LEGALLO, 2003; LESAGE e PACE, 2009). De qualquer maneira, é necessário detectar se, de fato, as observações vizinhas interferem de alguma forma.

O modelo em que o *spillovers* espaciais são observados na variável dependente denomina-se *Spatial Autoregressive Model* (SAR). Anselin e Bera (1998) atentam ao fato de incluir variáveis independentes no modelo sem que haja interferência espacial, o qual intitulou de *mixed regressive-spatial autoregressive model*. A forma da função é apresentada a seguir:

$$y = (\mathbf{I} - \rho\mathbf{W})^{-1}(\mathbf{X}\beta + u) \quad (1.2)$$

O modelo possui “n” observações e “k” parâmetros, onde “y” é um vetor de dimensão $n \times 1$, \mathbf{W} é a matriz de vizinhança de ordem $n \times n$, \mathbf{X} é uma matriz $n \times k$, “ β ” é o vetor de parâmetros de dimensão $k \times 1$, “u” é um vetor $n \times 1$ e “ ρ ” o coeficiente espacial autorregressivo que mede o grau de dependência espacial. O termo $\rho\mathbf{W}y$ é comumente chamado de *lag* espacial e compreende a influência das características representadas por “y” entre os vizinhos. Caso o coeficiente “ ρ ” não seja significativo, temos o modelo clássico de MQO.

Pode ser que a dependência espacial não se limite apenas em termos da variável dependente, dependendo do processo gerador de dados. Além do *lag* espacial do regressando, é possível que existam simultaneamente *spillovers* espaciais originados em regressores omitidos no termo de erro e que devam ser explicitados na equação. O modelo que contém dependência espacial em ambos é chamado de *Spatial Durbin Model* (SDM), como mostra a seguinte equação:

$$y = (\mathbf{I} - \rho\mathbf{W})^{-1}(\mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + u) \quad (1.3)$$

Se “ $\gamma=0$ ”, então não há *feedback* espacial do regressor para observações vizinhas, e o modelo se equivale a um SAR. Cabe ressaltar que tanto no SAR quanto no SDM o termo de erro segue uma distribuição Normal e apresenta matriz de covariância homoscedástica, isto é: $u \sim N(0, \sigma^2\mathbf{I})$ (LESAGE e PACE, 2009).

Quando o processo gerador de dados contém dependência espacial apenas no termo de erro tem-se o *Spatial Error Model* (SEM). Neste caso, a dificuldade de interpretação dos dados aumenta porque a dependência está em fatores que não são percebidos pelo econométrico, o que não acontece nos dois modelos anteriores, por exemplo. O modelo de regressão pode ser escrito com a seguinte notação:

$$y = \mathbf{X}\beta + \varepsilon, \quad (1.4)$$

onde “ ε ” é o termo de erro com processo espacial autoregressivo (nx1). Dentro dessa distúrbância encontra-se um erro totalmente aleatório, a saber:

$$\varepsilon = \lambda \mathbf{W}\varepsilon + u, \quad (1.5)$$

onde \mathbf{W} é a matriz de pesos (nxn), “ λ ” é o coeficiente espacial e “ u ” o erro aleatório (nx1). Substituindo (1.4) em (1.5) temos a forma reduzida:

$$y = \mathbf{X}\beta + (\mathbf{I} - \lambda \mathbf{W})^{-1}u \quad (1.6)$$

Outro modelo importante a se considerar é o SAC, que implementa dependência espacial no termo de erro e na variável dependente, sendo uma mistura de SAR com SEM. Sua forma funcional segue o seguinte esquema:

$$y = \rho \mathbf{W}_1 y + \mathbf{X}\beta + \varepsilon, \text{ onde } \varepsilon = \lambda \mathbf{W}_2 \varepsilon + u \quad (1.7)$$

Fazendo as devidas substituições e desenvolvendo a equação, tem-se:

$$y = (\mathbf{I} - \rho \mathbf{W}_1)^{-1} \mathbf{X}\beta + (\mathbf{I} - \lambda \mathbf{W}_2)^{-1} (\mathbf{I} - \rho \mathbf{W}_1)^{-1} u \quad (1.8)$$

Lesage e Pace (2009) atentam que a matriz \mathbf{W}_1 pode ser igual à matriz \mathbf{W}_2 , passando a se chamar de SARMA (*Spatial Autoregressive Moving Average Model*).

Nos modelos onde é encontrado algum tipo de dependência espacial a estimação por MQO gera estimadores viesados. Para tratar dessa questão, Lee (apud LESAGE e PACE, 2009) sugere a estimação por máxima verossimilhança. Uma hipótese importante para gerar a função de verossimilhança é a de que o termo de erro segue uma Distribuição Normal com *noise variance* constante (homoscedástico). Com a estimação por Máxima Verossimilhança, os parâmetros estimados são assintoticamente consistentes e torna-se possível realizar inferência estatística.

Para sugerir qual a especificação mais apropriada para representar os dados, o teste do Multiplicador de Lagrange (LM) é bastante apropriado. É o teste assintótico mais utilizado na literatura de econometria espacial por necessitar da estimação do modelo apenas sob a hipótese nula (ANSELIN et al, 2000). Ele pode ser realizado em sua versão tradicional ou robusta, tanto para o modelo que incorpora lag-espacial na variável dependente quanto no erro. No primeiro caso, temos as formas:

$$LM_{Lag} = \frac{\left(\frac{e'Wy}{(e'e)/n} \right)^2}{\frac{(WX\hat{\beta})'M(WX\hat{\beta})}{(e'e)/n} + tr(W'W+W^2)}, \quad (1.9)$$

$$RLM_{Lag} = \frac{\left[\frac{e'Wy}{(e'e)/n} - \frac{e'We}{(e'e)/n} \right]^2}{\frac{(WX\hat{\beta})'M(WX\hat{\beta})}{(e'e)/n} + tr}, \quad (1.10)$$

onde “ tr ” é o traço da matriz $M = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ e “ $\hat{\beta}$ ” é o estimador de MQO. O teste com o modelo de erro espacial, nas versões tradicional e robusta, assume as seguintes funções:

$$LM_{err} = \frac{\left(\frac{e'W_e}{(e'e)/n}\right)^2}{tr(W'W+W^2)}, \quad (1.11)$$

$$RLM_{err} = \frac{\left[\frac{e'W_e}{(e'e)/n} - tr\left(\frac{(WX\hat{\beta})'M(WX\beta)}{(e'e)/n} + tr\right)^{-1} \frac{e'Wy}{(e'e)/n}\right]^2}{tr - tr^2\left(\frac{(WX\hat{\beta})'M(WX\beta)}{(e'e)/n} + tr\right)^{-1}}, \quad (1.12)$$

As estatísticas do teste LM seguem uma distribuição Qui-Quadrado com apenas 1 grau de liberdade. Deve-se atentar que as versões robustas serão válidas apenas se as versões padrões forem significantes. Para decidir qual das opções é a melhor, deve-se observar qual teste LM tem o menor p-valor – ou escolher a alternativa com maior estatística LM. Se $RML > LM$ e for significativa, então deve-se selecionar a alternativa com maior RLM.

Em um modelo clássico de regressão linear os coeficientes das variáveis independentes indicam a correlação parcial com o regressando, sendo a derivada parcial da var. dependente em relação à independente. Quando se verifica que o processo gerador de dados incorpora a influência de vizinhos, a interpretação dos estimadores torna-se mais rebuscada, já que a variável explicativa e/ou dependente defasada é capaz de afetar a sua própria região e também a seus vizinhos geográficos (LESAGE e PACE, 2009).

Os modelos espaciais de regressão são capazes de ampliar o conjunto de informações para incluir outras oriundas de observações vizinhas, e os impactos estimados podem ser traduzidos pelos efeitos diretos, indiretos e totais. É possível gerar esses efeitos por meio dos modelos SAR, SDM e SAC - pois o *spillover* é observável –, trazendo assim uma interpretação rica e intuitiva. Como discutido anteriormente, a interpretação de dependência espacial nos modelos SEM é muito difícil e cabe acrescentar que não é possível obter essas medidas de impacto. Formalizando as três medidas citadas, temos:

Impacto Direto Médio – medida sumária que representa uma média dos efeitos de uma região “i” sobre ela mesma. Matematicamente é expressa como o traço da diagonal principal de $S_r(W)$, ou seja:

$$\overline{M}(r)_{direct} = n^{-1} tr(S_r(W)), \quad (1.13)$$

onde $S_r(W) = (I_n - \rho W)^{-1} I_n \beta_r$, “n” é o número de observações, “tr” é o traço da matriz e $\overline{M}(r)_{direct}$ é o efeito direto médio.

Impacto Total Médio – esta medida captura o impacto médio de todas as regiões incorporadas na matriz de vizinhança sobre uma determinada região “i”, incluindo o seu próprio efeito. Pode ser expressa da seguinte maneira:

$$\overline{M}(r)_{total} = n^{-1} t_n' S_r(W) t_n, \quad (1.14)$$

onde t_n é um vetor coluna unitário e $\overline{M}(r)_{total}$ é o efeito total médio.

Impacto Indireto Médio – Mede a influência das observações vizinhas sobre a região “i”, sendo obtida pela diferença entre os efeitos Total e Direto, ou seja:

$$\overline{M}(r)_{indirect} = \overline{M}(r)_{total} - \overline{M}(r)_{direct}, \quad (1.15)$$

ANÁLISE DOS RESULTADOS

Os resultados das estimações estão apresentados nesta seção e o objetivo é demonstrar de que forma o ambiente – juntamente com os fatores socioeconômicos - se correlacionam com os homicídios, procurando captar algum efeito de *spillover* das variáveis consideradas e também os efeitos diretos e indiretos. Tem-se por hipótese que os bairros que apresentam uma menor concentração/proporção de estabelecimentos residenciais, em média, estão correlacionados com maior incidência de homicídios, dado que um ambiente com poucas testemunhas pode favorecer a ação do criminoso (BRANTINGHAM e BRANTINGHAM, 1995). Pelos dados, observa-se que a ocorrência de homicídios nos turnos da noite e madrugada é maior. Esse é o momento em que as ruas ficam mais desertas e há menos vigilantes e testemunhas - baixos *law* e *social enforcement*⁶ -, o que pode facilitar a ação do criminoso especialmente nos ambientes onde as pessoas não residem. Por outro lado, pode ser que o fato de ocorrer muitos homicídios em determinadas áreas do bairro leva as pessoas a quererem residir em outro local, dentro do mesmo bairro ou até mesmo em outro. Os dados observados compreendem o período de 2010, então, no curto prazo, é plausível considerar que não há mudança⁷ do tipo de estabelecimento nesse período, o que corrobora para argumentar certa exogeneidade das variáveis ambientais. A oferta de estabelecimentos é dada e, nesse prazo de um ano, assume-se que um estabelecimento tido como residencial, por exemplo, permanece destinado à moradia. Por esse prisma, a causalidade reversa entre ambiente e criminalidade não seria um impedimento e as variáveis socioeconômicas entrariam como controles. Torna-se complicado, então, controlar outras possíveis variáveis (não-espaciais) que estariam omitidas no termo de erro, o que enviesaria os estimadores. Como a endogeneidade mostra-se um problema difícil de se contornar, a busca pelas correlações é importante para averiguar o sentido da relação entre as variáveis, se está conforme a abordagem da teoria criminológica que enfatiza a importância do ambiente. Outro ponto essencial é que a forma como as variáveis ambientais foram construídas procura captar a ideia do *law* e *social enforcements*.

Foram testados 8 modelos empíricos que incorporam as características ambientais e socioeconômicas como variáveis explicativas e a taxa de homicídio como variável dependente. Foram utilizadas técnicas estatísticas para assegurar a escolha dos melhores modelos que representam a relação entre as variáveis⁸ para implementar a técnica de econometria espacial, visto que não foi implementado explicitamente um modelo teórico, apesar de que a inclusão das variáveis e o que elas nos informam estejam consoantes com a base teórica criminológica.

Dentre os modelos propostos, escolhemos 3 deles - modelos (5), (6) e (7) - para averiguar se há presença e algum potencial efeito espacial nos dados. Para realizar o Teste do Multiplicador de Lagrange foi necessário inserir a matriz de vizinhança no modelo a se estimar, sendo testadas as matrizes tipo Queen, inverso da distância e a *k-nearest neighbour*. A que melhor se ajustou aos dados foi a matriz de distância e foi utilizada para realizar as estimações, posteriormente. A Tabela 4 mostra os resultados das estatísticas de teste e o seu respectivo *p*-valor para o teste LM. Para obtenção das estatísticas, apenas a estimação sobre os modelos de MQO especificados na Tabela 3 foi necessária. De acordo com os resultados, os três modelos averiguados possuem algum tipo de dependência espacial por meio de *lag*-

⁶ O termo *law enforcement*, no contexto de crime, indica qualquer esforço por vias legais em coibir a atividade ilícita. O *social enforcement* representa um mecanismo social capaz de restringir qualquer tipo de ação passível de reprovação, como as normas sociais e o capital social, por exemplo.

⁷ Pode até haver mudança do tipo de estabelecimento, mas ela não é significativa dado que é necessário um tempo considerável para transformar um local residencial num comercial, por exemplo - o contrário também é válido.

⁸ Todos os modelos formulados e testes encontram-se no apêndice.

espacial, podendo encontrar algum efeito de *spillover* tanto em termos de criminalidade para bairros vizinhos quanto dos fatores socioeconômicos.

Tabela 4 - Teste do Multiplicador de Lagrange

	Estatística de Teste			<i>p</i> -valor		
	(5)	(6)	(7)	(5)	(6)	(7)
LMerr	1.1612	0.6606	0.9910	0.281	0.416	0.319
LMlag	3.0463	3.440	3.656	0.081	0.063	0.055
RLMerr	0.2145	0.8731	0.505	0.643	0.350	0.477
RLMlag	2.0995	3.653	3.1702	0.196	0.056	0.075

Os termos entre parênteses representam o *p*-valor. Os *p*-valores em negrito indicam significância estatística.

Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE, Censo Demográfico e CNEFE (IBGE).

Partindo da indicação de dependência espacial nos modelos por meio do teste LM, foram realizadas as estimações por máxima verossimilhança incorporando a matriz de distância nos 3 modelos. As hipóteses de normalidade e homoscedasticidade dos erros são importantes para garantir que a estimação gere estimadores não-viesados, e apenas a de normalidade para garantir a consistência. A Tabela 5 indica as estatísticas e valores de probabilidade dos testes *Koenker* e *Jarque-Bera*, que averiguam presença de heteroscedasticidade e não-normalidade, respectivamente. A heteroscedasticidade parece ser um problema apenas no modelo (5), enquanto que a hipótese de normalidade dos erros é violada nos três modelos propostos. Os gráficos 6.1, 6.2 e 6.3 mostram a plotagem dos resíduos dos modelos e, apesar de serem assimétricos, os resíduos dos modelos (5) e (6) assemelham-se mais com uma curva Normal que os do modelo (7). Como a amostra é pequena (94 observações), assumiremos a Lei dos Grandes Números que garante que a distribuição dos erros converge em probabilidade para uma Distribuição Normal quando a amostra tende ao infinito para poder realizar as inferências.

As estimações dos modelos SAR, SAC e SDM são apresentadas na Tabela 6. Os três modelos espaciais incorporam a variável dependente defasada espacialmente, embora o SAC e o SDM incluam a dependência nos resíduos e nos regressores, respectivamente. Vale salientar que o vetor “ β ” não pode ser interpretado da mesma maneira que em regressão clássica porque se deve considerar o peso das observações imposto pela matriz de vizinhança, ponto que será discutido na subseção posterior. Mesmo que a interpretação seja feita através das análises dos efeitos espaciais, a significância estatística de todos os coeficientes estimados por máxima-verossimilhança é importante porque os mesmos são argumentos na função dos efeitos Direto e Total.

Gráfico 6.1 – Distribuição dos Resíduos do Modelo (5)

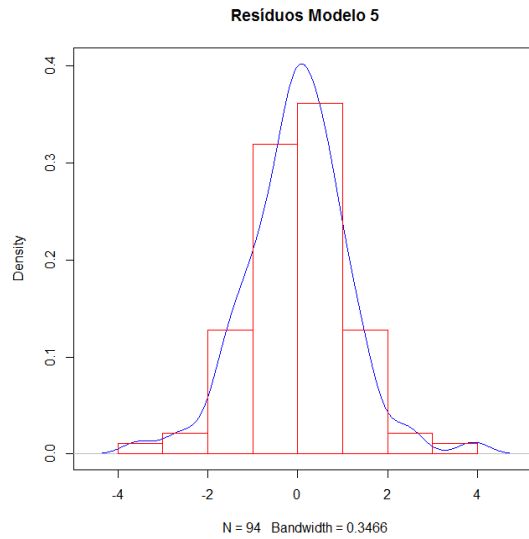


Gráfico 6.2 – Distribuição dos Resíduos do Modelo (6)

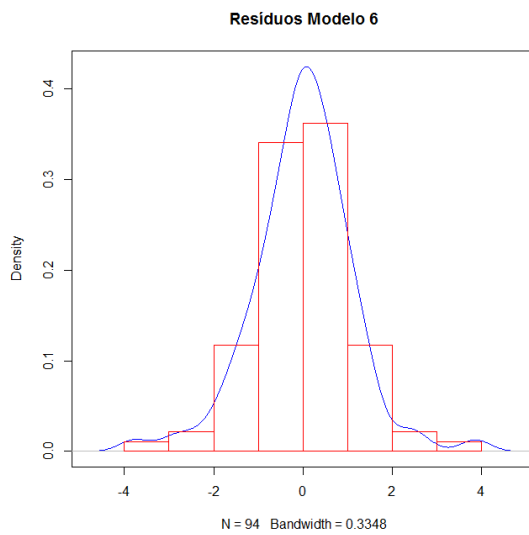


Gráfico 6.3 – Distribuição dos Resíduos do Modelo (7)

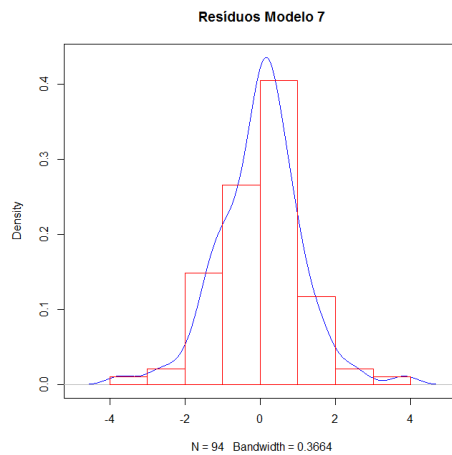


Tabela 5: Análise de Heteroscedasticidade e Normalidade dos Resíduos

Koenker	p -valor	Jarque-Bera	p -valor
---------	------------	-------------	------------

Modelo 5	16.769	0.018	9.05	0.010
Modelo 6	11.05	0.136	16.267	0.0002
Modelo 7	10.57	0.158	11.951	0.002

Os termos entre parênteses representam o *p*-valor. Os *p*-valores em negrito indicam significância estatística.

Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE, Censo Demográfico e CNEFE (IBGE).

Em nenhum dos modelos SDM o parâmetro *rhô* foi significativo, indicando que não há *spillover* espacial por essa modelagem. Nos modelos SAC, a dependência espacial na criminalidade é observada nos modelos (6) e (7), embora o coeficiente *lambda* – que capta a influência espacial nos resíduos – não seja significativa. Os modelos com os menores resultados do critério AIC (modelos (6) e (7)) incluem apenas o *lag* da variável dependente, com parâmetro *rhô* negativo e significativo, resultado já encontrado anteriormente por Menezes et al (2011) - embora incluindo um conjunto de variáveis explicativas diferente -, denotando que a incidência de criminalidade nos bairros vizinhos está associada a uma redução da criminalidade num determinado bairro, de modo geral. Entre os dois modelos mencionados, o modelo (6) apresenta melhor ajuste em comparação com os dois restantes por apresentar a menor estatística pelo critério AIC.

Sem a obtenção dos efeitos sumários, não é possível fazer uma análise dos estimadores, tendo em vista que se deve considerar a informação incorporada pelas observações vizinhas. A análise das medidas sumárias foi realizada sob o modelo (6) por ter apresentado a melhor configuração de variáveis. Com a estimativa dos efeitos marginais é possível compreender de que forma uma característica de um determinado vizinho pode gerar um efeito de *feedback* num determinado bairro observado.

Deve-se ressaltar que esse tratamento corrige viés de variável omitida espacialmente, não permitindo tratar viés de outras variáveis. Mesmo existindo o viés por conta da omissão de variáveis explicativas, o viés dos estimadores é reduzido com o tratamento econométrico espacial. O problema da causalidade reversa é amenizado no curto prazo dado que não há alteração da forma que o estabelecimento é utilizado. As pessoas podem até optar por residir em locais onde se sintam mais seguras, mas o mesmo estabelecimento permanece destinado a abrigar indivíduos, e outros passarão a morar nesse local. O fato de o crime ser menos frequente numa determinada região não provocaria uma mudança no tipo de estabelecimento com o intuito de melhorar o *law* e *social enforcements*, dado que as pessoas irão arbitrar em residir ou não no local. Portanto, o uso dos efeitos diretos, indiretos e totais passam a ser pertinentes mesmo com o viés gerado por variáveis omitidas.

Tabela 6 - Estimações dos modelos SAR, SAC e SDM

	SAR	SAC	SD M	SAR	SAC	SDM	SAR	SAC	SDM
	(5)	(5)	(5)	(6)	(6)	(6)	(7)	(7)	(7)
dens10	0.053 (0.015)	0.053 (0.016)	0.052 (0.016)	0.063 (0.054)	0.064 (0.052)	0.060 (0.066)			

Indens))))		1.443	1.44	1.477
10							(0.000)	8	(0.000)
)	(0.000)	
lnpop1	2.001	2.001	1.906	1.910	1.942	1.755	0.498	0.50	1.773
0	(0.010)	(0.012)	(0.030)	(0.003)	(0.007)	(0.020)	(0.022)	6	(0.038)
))))))	(0.020)	
gini	4.195	4.199	3.8531	3.784	3.877	3.240	3.403	3.44	2.4094
		7			6	8	5	4	
	(0.016)	(0.016)	(0.0445)	(0.023)	(0.018)	(0.070)	(0.04)	(0.041)	(0.186)
)))))			
lnrenda	-3.024	-3.017	-3.017	-3.406	-3.400	-3.384	-0.78	-	-2.648
	(0.002)	(0.004)	(0.009)	(0.000)	(0.000)	(0.000)	(0.000)	0.77	(0.005)
))))))	1	
								(0.002)	
lnnaores	0.537	0.529	0.548				0.459	0.45	1.718
s	(0.078)	(0.090)	(0.067)				(0.001)	8	(0.001)
))	(0.001)	
naores				5.805	5.746	5.481			
				4	9	2			
				(0.000)	(0.000)	(0.000)			
)))			
prop.nres				0.068	0.078	0.100	0.001	0.00	0.598
es								1	
				(0.576)	(0.518)	(0.423)	(0.000)	(0.000)	(0.000)
))))		
lnprop.nres	2.946	2.961	2.954						
	(0.003)	(0.004)	(0.003)						
)							
prop.res				-0.060	-0.061	-0.061			
s									
				(0.031)	(0.029)	(0.028)			
)	1))			
lnprop.res	-1.4113	-	-1.3949						
	(0.001)	1.412	(0.001)						
		5							
		(0.001)							
)							
lnprop.tot							-2.123	-	-2.1693
							(0.007)	2.12	(0.002)
)	7	
								(0.000)	

								07)	
lag.den		0.046				-0.019			-2.387
s		(0.718)				(0.96)			(0.43)
lag.lnp		-1.604				-1.559			-1.626
op		(0.4109)				(0.414)			(0.27)
))			
lag.gini		-6.1704				-			3.044
		(0.391)				8.655			(0.08)
						8			
						(0.197)			
)			
lag.ren		0.985				-0.835			0.572
da		(0.923)				(0.927)			(0.821)
)			
lag.lnn		2.891				1.160			1.420
aores		(0.2036)				4			(0.332)
)				(0.868)			
)			
lag.lnpr		-1.8965				-0.117			-0.108
op.nres		(0.1430)				(0.402)			(0.348)
))			
lag.lnpr		0.675				-0.048			-0.260
op.res		(0.971)				(0.883)			(0.999)
)			
constan	7.4530	7.421	7.149	1.915	1.915	3.289	6.455	6.36	6.360
te		0		8	6		8	91	
	(0.000)	(0.000)	(0.001)	(0.33)	(0.329)	(0.063)	(0.003)	(0.0	(0.005)
))))	03)	
Rho	-0.4107	-0.417	-0.412	-0.445	-0.516	-0.350	-0.446	-	-0.385
	(0.055)	(0.161)	(0.136)	(0.034)	(0.035)	(0.21)	(0.029)	0.47	(0.181)
))	4))	4	
								(0.0	
								60)	
Lambda		0.029			0.196			0.10	
		4			(0.573)			5	
		(0.948))			(0.7	
)						8)	
AIC	304.9	306.9	311.52	302.4	304.2	310.1	302.7	304.	307.98
		3		5	8	9		79	

Os termos entre parênteses representam o *p*-valor. Os *p*-valores em negrito indicam significância estatística.

Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE, Censo Demográfico e CNEFE (IBGE).

A Tabela 7 mostra os resultados dos efeitos e pode-se observar que os sinais são os mesmos obtidos com as regressões multivariadas, corroborando para assegurar o sentido dos mesmos. Todas as variáveis socioeconômicas e ambientais - com exceção da densidade de estabelecimentos não-residenciais que não apresenta nenhum efeito significativo - apresentam efeitos totais significativos e apenas os efeitos indiretos da densidade demográfica e de estabelecimentos residenciais são nulos. A significância dos efeitos totais é importante porque nenhum efeito que seja nulo invalida algum efeito significativo. Como a variável dependente é a taxa de homicídio linearizada, os efeitos médios são interpretados como elasticidades e semi-elasticidades, em caso de regressor linearizado ou em nível, respectivamente.

O fato de um bairro ter um grande adensamento populacional gera um efeito médio direto positivo sobre sua taxa de homicídio, mesmo que a densidade demográfica dos bairros vizinhos não afete a criminalidade do mesmo bairro observado (efeito indireto nulo). A quantidade de pessoas residentes num bairro também tem, em média, efeito direto positivo nos homicídios, visto que existem mais potenciais alvos aos olhos do agressor. Já o efeito indireto tem sentido contrário, representando que um maior nível populacional num bairro vizinho está associado a um menor nível de criminalidade num determinado bairro observado, ocorrendo uma espécie de migração de agressores para as regiões mais populosas. Nos bairros mais desiguais as taxas de homicídios são maiores, em média, e os bairros vizinhos que apresentam alto nível de desigualdade impactam negativamente na criminalidade do bairro observado.

Em relação às variáveis ambientais, os efeitos da proporção de estabelecimentos não-residenciais são fortemente significantes, conforme a Tabela 7. Quanto maior a proporção de locais destinados a fins que não sejam a residência particular em um bairro maior é o seu nível de criminalidade, sendo que o efeito indireto tem sentido contrário. Um maior adensamento de estabelecimentos residenciais num bairro parece estar associado a uma criminalidade menos intensa e o efeito de *feedback* é inexistente, ou seja, a concentração de residências nos bairros vizinhos não impacta na taxa de homicídio do bairro observado.

Tabela 6.5 - Resultados dos Efeitos Direto, Indireto e Total do Modelo SAR (6)

	Efeito Direto	Efeito Indireto	Efeito Total
dens10	0.064 (0.053)	-0.019 (0.152)	0.044 (0.070)
lnpop10	1.939 (0.002)	-0.596 (0.032)	1.342 (0.000)
gini	3.841 (0.024)	-1.180 (0.112)	2.660 (0.0391)
lnrenda	-3.447 (0.000)	1.059 (0.029)	-2.387 (0.0118)
Naores	5.886 (0.000)	-1.809 (0.047)	4.077 (0.001)
prop.nres	0.068 (0.580)	-0.021 (0.608)	0.047 (0.590)
prop.res	-0.060 (0.033)	0.018 (0.131)	-0.042 (0.044)

Os termos entre parênteses representam o *p*-valor. Os *p*-valores em negrito indicam significância estatística.

Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE, Censo Demográfico e CNEFE (IBGE).

CONCLUSÃO

Este trabalho teve por objetivo investigar a relação entre as características do ambiente e a criminalidade na cidade do Recife, tendo como *proxy* a taxa de homicídio, em vista à escassez de informações empíricas a esse respeito e em nível tão desagregado quanto os bairros. O estudo foi inovador por ter incluído informações sobre o ambiente com o auxílio de um banco de dados bastante atual e ainda pouco explorado pelos pesquisadores. Além de estudar tais características ambientais, foi analisado como os fatores socioeconômicos estão associados com a criminalidade em conjunto com as variáveis ambientais construídas.

O foco foi averiguar o sentido das correlações entre os fatores e a taxa de homicídio do Recife, já que problemas advindos de viés de variável omitida de qualquer natureza são comuns em estudos econômicos e exigem técnicas estatísticas diferentes para tratar a questão (BIDDERMAN et al, 2010; MUSSE, 2012; LEIGH e NEIL, 2010). Como foi discutido, a causalção reversa entre ambiente e crime pode ser amenizada devido à não alteração do tipo de estabelecimento dentro do período analisado, permitindo avaliar o ambiente como um preditor para a criminalidade dos bairros. Dentro desse contexto, para contornar o viés de variável espacial omitida, foi proposta a metodologia de econometria espacial a fim de investigar os efeitos das observações vizinhas.

Os resultados das regressões obtidos com o modelo SAR demonstram concordância com os estudos de criminologia que apontam o espaço como fator extremamente relevante para explicar o crime (BRANTINGHAM e BRANTINGHAM, 1995; COHEN e FELSON, 1979). Na cidade do Recife observa-se que nos bairros onde a proporção de não-residências em relação ao total de domicílios é alta ocorrem mais assassinatos, e uma possível explicação para isso é que nesses locais o *law* e o *social enforcements* são mais baixos e viabilizam os crimes. O crime violento letal e intencional é, por exemplo, muito mais notório que um furto ou roubo e, por isso, os locais onde o agressor é menos visado é preferível. Por bairro, a maioria dos homicídios ocorre no período da noite e madrugada, em sua maior parte em logradouros públicos, e os locais caracterizados por não haver residências particulares têm menos transeuntes nesses períodos do dia e, conseqüentemente, menos testemunhas.

A significância da variável que indica o adensamento de residências nos bairros corrobora com essa conjectura, visto que existem mais barreiras à prática do crime nos lugares onde as pessoas moram mais próximas uma das outras, dado o trânsito de indivíduos e maior número de testemunhas. Por conseqüência, se um bairro vizinho possui um ambiente mais favorável e condizente com o cenário discutido, o agressor se torna mais apto a migrar para essas regiões e cometer o homicídio. A concentração (densidade) de estabelecimentos não-residenciais parece não importar para a ocorrência de homicídios, provavelmente porque o agressor escolhe pontos estratégicos no ambiente.

Atentando para as correlações entre taxa de homicídio e as variáveis econômicas, obteve-se um resultado intrigante em relação à densidade demográfica, que contraria o encontrado por Menezes et al. (2011). O efeito direto da densidade é positivo na criminalidade, indicando que se encontram mais crimes em regiões onde a concentração de pessoas é maior num determinado bairro, embora esse efeito não tenha *spillover* para regiões vizinhas. Pelos mapas Kernel, observa-se que existe uma correlação positiva muito forte entre

áreas residenciais e densidade demográfica, quando era esperado que a última mantivesse o mesmo sentido (negativo) em relação à taxa de homicídio. Por outro lado, Glaeser e Sacerdote (1999) demonstraram que nos grandes centros urbanos a prática do crime é facilitada devido a menor probabilidade de o agressor ser identificado, sendo uma possível justificativa para o sinal encontrado para os coeficientes da densidade demográfica e também do nível populacional.

O alto nível de desigualdade de renda está associado a uma alta incidência de crime nos bairros do Recife e corrobora para reforçar a ideia do baixo nível de *social enforcement* dos bairros, de modo geral. Por outro lado, a elevada renda *per capita* do bairro correlaciona-se negativamente com a taxa de homicídio. De fato, os homicídios são mais recorrentes nos bairros mais pobres da cidade e, o fato de os bairros serem bastante heterogêneos (residências de classe média-alta vizinhas a comunidades carentes dentro de um mesmo bairro), favorece o resultado encontrado para o coeficiente do índice de *gini* e está consoante com a literatura (FANJZYLBER et al., 1998; FANJZYLBER et al., 2000; KELLY, 2000; MENEZES et al., 2011).

Os resultados deste estudo trazem importante conteúdo a respeito do perfil da cidade do Recife no âmbito socioeconômico e ambiental de maneira a associá-los com a ocorrência de homicídios. É possível reduzir a criminalidade observando o ambiente e não apenas os mecanismos de incentivo dos criminosos, já que a realização de um ato ilícito pode ser facilitada de acordo com as características do local. Dado que a grande maioria dos crimes envolvendo morte ocorrem em logradouro público (ambiente aberto), uma forma de se investir na redução e controle da criminalidade é aumentando o *law enforcement* - vigilância por parte da polícia civil nas ruas não apenas em áreas residenciais; câmeras filmadoras de alta resolução que auxiliam na identificação dos criminosos nos ambientes mencionados -, principalmente nos lugares onde as pessoas não residem e onde menos transitam. Para aumentar a eficiência do investimento em *law enforcement* e determinar estrategicamente os locais a receber maior reforço, poderia ser elaborado um estudo para detectar as maiores probabilidades de ocorrência de assassinatos condicionadas ao período do dia, e também do dia da semana. O efeito de *spillover* da criminalidade entre os bairros denota que o agressor busca o local mais favorável ao crime, então privilegiar a estratégia de implementação do *enforcement* em apenas alguns bairros pode não ser a mais eficiente. Portanto, esta medida seria um complemento às políticas de combate ao crime já vigentes na cidade do Recife.

Por fim, este trabalho teve por objetivo principal identificar a importância do ambiente para como fator associado à criminalidade. É certo que o processo de modelagem empírica é extremamente complexo, dado que o econometrista não é capaz de observar uma série de fatores, além de que podem existir outros modelos que se ajustam melhor aos dados. Uma sugestão para trabalhos futuros é incluir outras variáveis socioeconômicas e utilizar a Regressão Beta com econometria espacial, visto que a variável dependente é estudada em termos de taxa.

APÊNDICE

Esta seção apresenta algumas configurações de modelos em que são inseridas as variáveis de ambiente e socioeconômicas descritas na Tabela 3. Seguindo a ideia de Brantingham e Brantingham (1995) de que o ambiente carrega circunstâncias que propiciam a prática do crime, procurou-se averiguar se a relação crime/ambiente ocorre e em qual direção. Foram oito modelos testados no total, ocorrendo diferenças entre eles em termos da linearidade da variável dependente e independente e também das variáveis de ambiente. Em cada modelo foram consideradas 3 variáveis ambientais, não sendo incluídas as quatro simultaneamente para contornar o problema de multicolinearidade perfeita, o que tornaria

algum coeficiente indeterminado matematicamente. Os dois primeiros modelos da Tabela 3 possuem todas as variáveis sem sofrer linearização (lin-lin), e o intuito é comparar com os modelos log-log (3) e (4) através do teste de especificação de Box-Cox. Em ambos os testes a hipótese nula é rejeitada, que postula que um dos modelos é superior ao outro, e pelo fato de o valor absoluto do numerador da estatística de teste ser maior que o denominador, o modelo com as variáveis loglinearizadas é mais representativo. O teste RESET e o critério de AIC também corroboram com a decisão de que os modelos (1) e (2) não são representativos dos dados, em comparação com os modelos (3) e (4).

Tendo em vista que a forma não linearizada dos modelos (1) e (2) das variáveis não se ajustou tão bem aos dados quanto à linearizada - e pelos resultados obtidos do teste RESET e critério de AIC -, partiu-se para a configuração de mais quatro modelos, agora realizando uma permutação entre variáveis linearizadas e outras em nível, embora todos apresentem a renda e a população em log neperiano. Em todos os modelos testados com a variável dependente log-linearizada, os coeficientes das variáveis socioeconômicas apresentam o mesmo sinal e são estatisticamente significativos, embora os efeitos das mesmas variáveis sejam diferentes em cada modelo. Os coeficientes das variáveis de proporção de estabelecimentos não-residenciais e densidade de estabelecimentos não-residenciais possuem sinal positivo, tanto com as variáveis em nível quanto linearizadas. Apenas nos modelos (4) e (6) a variável que representa a densidade de estabelecimentos não-residenciais não apresenta coeficiente significativo. Tanto a correlação entre a taxa de homicídio com a densidade de estabelecimentos residenciais quanto com a densidade de estabelecimentos totais é negativa e significativa nos modelos propostos.

Tabela 3 – Estimação por MQO dos modelos multivariados

	Variável dependente é a taxa de homicídio		Variável dependente é o logaritmo natural da taxa de homicídio					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
dens2010km ²	0.541 (0.006)	0.541 (0.006)			0.058 (0.015)	0.018 (0.057)		
lndens2010			1.33 (0.008)	0.98 (0.018)			1.47 (0.002)	
dlndenss								0.83 (0.04)
pop2010	0.062 (0.290)	0.062 (0.290)						
lnpop2010			0.47 (0.000)	0.53 (0.000)	2.130 (0.030)	2.001 (0.022)	0.53 (0.022)	0.48 (0.000)
gini	1.329 (0.002)	1.329 (0.002)			4.28 (0.0240)	3.98 (0.0295)	3.42 (0.064)	3.44 (0.08)
lngini			1.80 (0.0537)	1.56 (0.087)				
renda	0.383 (0.187)	0.383 (0.187)						
lnrenda			-0.88 (0.005)	-0.89 (0.000)	-2.906 (0.008)	-3.347 (0.001)	-0.77 (0.013)	-1.05 (0.000)

naores	0.069 (0.004)	0.069 (0.004)				5.96 (0.000))
Innaores			0.30 (0.0124)	0.22 (0.0478)	0.529 (0.10752)		0.45 (0.014)	0.14 (0.11)
prop.nres	1.593 (0.0019)	1.133 (0.0135)				0.050 (0.700)	0.002 (0.000)	
Inprop.nres			0.90 (0.021)	0.50 (0.152)	7.988e-01 (0.007)			0.77 (0.06)
prop.res		-0.456 (0.006)				-0.059 (0.049)		
Inprop.res				-1.58 (0.006)	-1.45 (0.000)			
prop.tot	-0.456 (0.019)							
Inprop.tot			-2.287 (0.017)				-2.15 (0.009)	-0.92 (0.01)
constante	-59.11 (0.032)	-59.11 (0.032)	6.37 (0.002)	5.16 (0.015)	15.87 (0.003)	0.282 (0.99)	4.58 (0.037)	5.47 (0.00)
R ² ajustado	0.61	0.615	0.51	0.535	0.539	0.548	0.545	0.49
AIC	1273.8	1273.86	312.27	307.25	306.486	304.72	305.16	314.7
Teste RESET	6 (0.00)	9 (0.00)	9 (0.122)	6 (0.114)	(0.86)	(0.86)	(0.51)	2 (0.32)
Teste Box-Cox	(0.00)	(0.00)						

Os termos entre parênteses representam o *p*-valor. Os *p*-valores em negrito indicam significância estatística

Fonte: Elaborado pelo Autor a partir das informações do INFOPOL SDS/PE, Censo Demográfico e CNEFE (IBGE).

O sinal positivo do coeficiente da taxa de crescimento populacional (dlndens) e da densidade demográfica (dens2010km2 e Indens2010) em todos os modelos corrobora com a teoria de Glaeser e Sacerdote (1999), onde o processo de urbanização traz implicações à criminalidade. A direção das correlações das outras variáveis socioeconômicas estão consoantes com Menezes et al (2011) exceto para a densidade demográfica, que nos modelos propostos influencia positivamente a taxa de homicídio.

Os modelos (6), (7) e (5) possuem os menores valores do AIC e maiores R² ajustado, respectivamente. O fato de a hipótese nula do teste RESET não ter sido rejeitada indica que a especificação do modelo é aceitável. Utilizando os dois critérios citados, selecionamos os 3 melhores ajustes de modelos para averiguar a existência de dependência espacial entre observações vizinhas e controlar o problema de viés de variável espacial omitida.

BIBLIOGRAFIA

- ALMEIDA, E. DE S.; HADDAD, E. A.; HEWINGS, G. J. D. **The Spatial Pattern of Crime in Minas Gerais: An Exploratory Analysis.** *Economia Aplicada*, v. 9(1), jan/mar 2005.
- ANDRADE, M. V.; LISBOA, M. B. **Desesperança de vida: Homicídio em Minas Gerais.** In: Henriques, R., editor, *Desigualdade e Pobreza no Brasil*, p. 347-384. IPEA, Rio de Janeiro, 2000.
- ANSELIN, L.; BERA, A. **Spatial dependence in linear regression models with an introduction to spatial econometrics.** In: *Handbook of applied economic statistics*, edited by Amman Ullah and David E.A. Giles. New York: Marcel Dekker. 1998.
- ANSELIN, L.; COHEN, J.; COOK, W. G.; TITA, G. **Spatial Analyses of Crime. Measurement and Analyses of Crime and Justice,** *Criminal Justice*, vol 4, 2000.
- BECKER, G. S. **Crime and punishment: An economic approach.** *The Journal of Political Economy*, v. 76(2), p. 169-217, 1968.
- BIDDERMAN, C.; DE MELLO, J. M. P.; SCHNEIDER, A. **Dry Laws and Homicides: Evidence from the São Paulo Metropolitan Area.** *The Economic Journal*, v. 120, issue 543, 2010.
- BRANTINGHAM, P.L.; BRANTINGHAM, P. J. **Nodes, paths and edges: considerations on environmental criminology.** *Journal of Environmental Psychology*, v. 13, p. 3-28, 1993b.
- BRANTINGHAM, P.J.; BRANTINGHAM, P.L., MOLUMBY, T. **Perceptions of crime in a dreadful enclosure Ohio.** *Journal of Science*, v. 77, p. 256-261, 1977.
- BRANTINGHAM, P.L.; BRANTINGHAM, P.J., SEAGRAVE, J. **Crime and fear at a Canadian university.** In: B.S. Fisher, J.J. Sloan (eds.), *Campus Crime: Legal Social, and Poficy Perspectives* Springfield (III), Charles C. Thomas, p. 123-155, 1995.
- BRANTINGHAM, P.; BRANTINGHAM, P. **Crime Generators and Crime Attractors.** *European Journal on Criminal Police and Research*, v. 3(3), p. 1-26, 1996.
- CLARKE, R.V.; MAYHEW, P. **Parking patterns and car theft risks: Policy Relevant Findings From The British Crime Survey.** In: R.V. Clarke (ed.), *Crime Prevention Studies*, Monsey (NY), Criminal Justice Press, v. 3, p. 91-107, 1994.
- COHEN. L.E.; M. FELSON. **Social Change and Crime Rate Trends: A Routine Activity Approach."** *American Sociological Review*, v. 44, p. 588-608, 1979.
- ECK, J.; SPELLMAN, W. **Thefts from vehicles in ship yard parking lots.** In: R.V. Clarke (ed.), *Situational Crime Prevention* New York, Harrow and Heston, p. 164-173, 1992.
- ERTUR, LEGALLO. **An Exploraty Spatial Data Analysis of European Regional Disparities, 1980-1995.** In: Fingleton, Growth. Springer, 2003.
- FAJNZYLBER, P.; LEDERMAN, D.; LOAYZA, N. **Determinants of Crime Rates in Latin America and the World.** Washington, DC: The World Bank, 1998.

FAJNZYLBER, P.; LEDERMAN, D.; LOAYZA, N. **Crime and Victimization: An Economic Perspective**. *Economia*, v. 1(1), p. 219-278, 2000.

FAJNZYLBER, P.; LEDERMAN, D.; LOAYZA, N. **'Inequality and Violent Crime'**, *Journal of Law and Economics*, v. 45, p. 1-39, 2002.

FLEMING, Z.; BRANTINGHAM, P.; BRANTINGHAM, P. **Exploring auto theft in British Columbia**. In: R.V. Clark (ed.), *Crime Prevention Studies*, v. 3 Monsey CNYJ, Criminal Justice Press, p. 47-90, 1994.

GLAESER, E. L.; SACERDOTE, B. **Why Is There More Crime in Cities?** *Journal of Political Economy*, University of Chicago Press, vol. 107(S6), p. S225-S258, dec. 1999.

HOFFMANN, R. **Mensuração da desigualdade e da pobreza no Brasil**. In: HENRIQUES, R. (org.). *Desigualdade e pobreza no Brasil*. Rio de Janeiro: Ipea, p. 81-107, 2000.

KELLY, M. **Inequality and Crime**. *The Review of Economics and Statistics*, November, v. 82(4), p. 530-539, 2000.

LAFREE, G. **A Summary and Review of Cross-National Comparative Studies of Homicide**, In: M. D. Smith and M. A. Zahn, eds, *Homicide: A Sourcebook of Social Research*, p. 125-45. Thousand Oaks, CA: Sage Publications, 1999.

LEIGH, A.; NEILL, C. **Do Gun Buybacks saves Lives? Evidence from panel data**. *Oxford Journals*, 2010.

LESAGE, J.; PACE, R. K. **Introduction to Spatial Econometrics**, CRC Press, 2009.

MENEZES, T.; SILVEIRA NETO, R. M.; MONTEIRO, S.; RATTON, J. **Spatial Correlation Between Homicide Rates And Inequality: Evidence From Urban Neighborhoods In A Brazilian City**. Paper supported by Pernambuco Research Foundation (FACEPE), 2011.

MESSNER, S. F. **Poverty, Inequality, and the Urban Homicide Rate: Some Unexpected Findings**, *Criminology*, v. 20, p. 103-14, 1982.

MESSNER, S. F.; ROSENFELD, R. **Political Restraint of the Market and Levels of Criminal Homicide: A Cross-National Application of Institutional-Anomie Theory**. *Social Forces*, v. 75, 393-416, 1997.

MUSSE, I. F. **O crime de drogas e a violência em São Paulo: Uma análise a partir da lei de drogas**, Fundação Getúlio Vargas; dissertação, 2012.

POYNER, B. **Situational crime prevention in two parking facilities**. In: R.V. Clarke (ed.), *Situational Crime Prevention* New York, Harrow and Heston, p. 174-184, 1992.

PRIDEMORE, W. A. **Poverty Matters: A Reassessment of the Inequality-Homicide Relationship in Cross-National Studies**. Published by Oxford University Press on behalf of the Centre for Crime and Justice Studies (ISTD), *BRIT. J. CRIMINOL.* v. 51, p. 739-772, 2011.

SCORZAFAVE, L. G., SOARES, M. K. **Income inequality and pecuniary crimes**, *Economics Letters*. Elsevier, 2009.

SOARES, R. R. Development, crime and punishment: accounting for the international differences in crime rates. Journal of Development Economics, v. 73, p. 155-184, 2004.

TYSZLER, M. Econometria Espacial: discutindo medidas para a matriz de ponderação espacial. Fundação Getúlio Vargas, 2006.

WILCOX, S. The Geography of Robbery; the Prevention and Control of Robbery, v. 3 Davis, Center of Administration of Justice, University of California, 1973.

WITTE, A. D. Estimanting the Economic Model of Crime With Individual Data. Published by John Wiley & Sons, Inc. The Quarterly Journal of Economics, feb. 1980.

WORTLEY, R.; MAZEROLLE, L. Environmental Criminology and Crime Analysis. Routledge, London and New York, 2008.