

"A FEA e a USP respeitam os direitos autorais deste trabalho. Nós acreditamos que a melhor proteção contra o uso ilegítimo deste texto é a publicação online. Além de preservar o conteúdo motiva-nos oferecer à sociedade o conhecimento produzido no âmbito da universidade pública e dar publicidade ao esforço do pesquisador. Entretanto, caso não seja do interesse do autor manter o documento online, pedimos compreensão em relação à iniciativa e o contato pelo e-mail bibfea@usp.br para que possamos tomar as providências cabíveis (remoção da tese ou dissertação da BDTD)."

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA

HOMICÍDIOS NA CIDADE DE SÃO PAULO
Uma Análise de Causalidade e Autocorrelação Espaço-Temporal

Alexandre Sartoris Neto

Orientadora: Profa. Dra. Vera Lucia Fava

DEDALUS - Acervo - FEA



20600004070

São Paulo
2000

Agradecimentos

A minha professora, orientadora e amiga, Vera Lucia Fava, que, como sempre, fez jus ao nome que tem.

Aos meus pais, pelo apoio.

À Silvia, de quem eu tenho o prazer de ser irmão.

À Sylvia, que não é minha irmã, mas eu gostaria que fosse.

À Gi, pela paciência.

Aos professores Paulo Picchetti, Maria Dolores Diaz e Heron do Carmo, pelas críticas e sugestões no exame de qualificação.

Aos professores Marcos Lisboa, Maria Sylvia Saes, Denisard Alves e, novamente, Heron do Carmo, pelas observações e comentários por ocasião da defesa.

Aos demais professores e colegas da FEA.

Ao CNPQ, pelo financiamento durante parte do trabalho.

ÍNDICE

Introdução	1
1. Abordagem Econômica do Crime	3
1.1 O crime de homicídio	20
1.2 Estudos no Brasil	25
1.3 O modelo de propagação	27
2. Causalidade	30
2.1 Definições de causalidade	31
2.2 Tipos de causalidade	35
3. Espaço e Tempo	38
3.1 Conceito de tempo	39
3.2 Conceito de espaço	41
3.3 Espaço na economia	42
3.4 Econometria Espacial	44
3.4.1 Estatística Espacial	45
3.4.2 Autocorrelação Espacial	48
3.4.3 Causalidade Espaço-Temporal	50
3.4.4. Os modelos STARMA	52
3.4.5. Outros desenvolvimentos em Econometria Espacial	55
4. Resultados Empíricos	56
4.1 Dados utilizados	56
4.2 Resultados em séries de tempo	61
4.3 Testes de autocorrelação espacial	69
4.4 Teste de causalidade no espaço-tempo	71
4.5 O modelo STARMA	72
4.5.1 Identificação	72
4.5.2 Estimação e diagnóstico	73
4.6 Simulações através do modelo STAR	76
5. Conclusões	81
Bibliografia	83

Introdução

Uma das questões urbanas mais relevantes nos últimos tempos é a da criminalidade.

Ela é, indubitavelmente, bastante complexa, e pode (e deve) ser abordada em diferentes aspectos, sendo o econômico o fundamento central para este trabalho.

Quando se fala em aspectos econômicos da atividade criminal, pode-se pensar em duas formas: variáveis econômicas (macroeconômicas principalmente) afetando o comportamento dos indivíduos. A outra, que tem em Becker(1968) e Ehrlich(1973) como dois de seus "fundadores", encara a escolha pelo crime como uma escolha racional, uma alocação que o indivíduo faz entre seu tempo, a exemplo da escolha entre trabalho e lazer, tradicional dos livros textos de microeconomia.

A abordagem utilizada neste trabalho é esta última. O crime a ser estudado é o homicídio. A hipótese é que existe uma relação causal entre o número de crimes em uma região (no caso deste trabalho, a divisão da cidade é feita pelas áreas de atuação dos distritos policiais, em dados mensais de 1995 e 1996) e o número de crimes na região vizinha no período anterior.

Esta relação de causalidade no espaço-tempo, que também poderíamos chamar de propagação é a questão que procura-se responder empiricamente. Para isso, utilizaremos os recursos da Econometria Espacial unidos aos recursos da Econometria de Séries de tempo.

No capítulo 1, trataremos da abordagem econômica do crime, especificamente no que se refere às motivações da atividade criminal, procurando-se definir uma função "oferta de crimes". Ao final do capítulo, será exposto o modelo teórico da relação de causalidade no espaço-tempo (propagação) para a atividade criminal.

A causalidade é o tema do capítulo 2. Neste capítulo discutiremos os conceitos de causalidade, culminando com o conceito de Granger(1969).

No capítulo 3 será feita uma breve apresentação das questões do espaço e tempo na ciência, em particular na ciência econômica. Também neste capítulo serão introduzidas a Estatística Espacial e a Econometria Espacial.

O capítulo 4 é dedicado aos resultados empíricos: primeiramente, testaremos um modelo puramente temporal, utilizando o número total de homicídios em cada mês, e depois um modelo puramente espacial (autocorrelação espacial), utilizando o total por distrito policial em cada ano. A seguir, a causalidade espaço-temporal será testada através de uma adaptação do teste de Causalidade de Granger à modelagem espacial, que chamaremos de teste de Causalidade de Granger Espaço-Temporal. Finalmente, será estimado um modelo do tipo STARMA (*space-time autoregressive moving average*) para a cidade, que servirá de base para simulações do que pode ocorrer nos próximos anos.

O capítulo 5 será destinado às conclusões finais.

1. Abordagem Econômica do Crime

O crime é, sem dúvida, um grande problema social e uma preocupação crescente em nossa sociedade. É uma questão que, por sua complexidade, demanda para seu entendimento mais completo, um tratamento multidisciplinar. Neste texto, porém, o foco será o aspecto econômico da atividade criminosa de um modo geral, e do crime de homicídio em particular, principalmente no que se refere às suas motivações.

Nesta abordagem, o crime é visto, via de regra, como um comportamento racional, isto é, uma escolha de uma pessoa ou grupo de pessoas de como utilizar seu tempo, entre atividades legais ou criminosas. Para fazer esta escolha, considera-se que os indivíduos têm conhecimento, não necessariamente perfeito, dos custos e benefícios associados a ela, ou seja, o crime é visto como uma atividade econômica qualquer. Esta abordagem segue a linha, entre outros, de Becker(1968), Ehrlich(1973 e 1996), Hellman e Alper(1997) e Block e Heineke(1975).

O estudo do crime, por este prisma, não é algo universalmente aceito. Horton e Leslie(1970, p.167), por exemplo, argumentando contra a relação de causa efeito entre maior punição e menor criminalidade, negam a idéia de racionalidade presente na decisão pela atividade criminal:

A equivocada fé na punição deve se basear na hipótese irrealista de que as pessoas conscientemente decidem ser ou não criminosas — que elas consideram uma carreira criminosa, racionalmente pesam seus perigos contra suas recompensas, e chegam a uma decisão baseada num cálculo de prazer e dor. Segue-se, supostamente, que se o elemento de dor é acrescido de severas punições, as pessoas sairão do crime para a legalidade. Um pouco de reflexão revela o absurdo desta noção.

Não se pretende, contudo, negligenciar a importância de outras causas de natureza sociológica, antropológica ou mesmo psicológica presentes na criminologia tradicional, mas é difícil imaginar, por exemplo, o

aumento de criminalidade em determinadas situações sem a presença de algum tipo de decisão racional. Nas palavras de Ehrlich(1973, p.522):

Uma dependência numa motivação única (resultado de circunstâncias familiares ou sociais excepcionais) como a grande explicação para o crime não traz, em geral, previsões a respeito do resultado de circunstâncias objetivas. Também não temos conhecimento de qualquer evidência empírica persuasiva registrada na literatura que apóie teorias usando esta abordagem. Nosso ponto de vista alternativo, ainda que não incompatível, é que mesmo que aqueles que violam certas leis difiram sistematicamente em vários aspectos daqueles que seguem estas mesmas leis, estes últimos, como os primeiros, respondem a incentivos. Mais do que recorrer a hipóteses a respeito de características pessoais únicas e condições sociais que afetem o respeito à lei, tendência à violência, preferência por risco ou preferência pelo crime de um modo geral, os primeiros podem ser separados por oportunidades mensuráveis e ver em que extensão o comportamento ilegal pode ser explicado pelo efeito das oportunidades dadas as preferências.

Esta abordagem aparece com destaque na literatura a partir do artigo de Becker(1968). Neste artigo foram enfocados vários aspectos da atividade criminal, como sua motivação, prevenção e repressão. No presente texto se dará particular atenção à primeira. Becker definiu uma função para o número de crimes cometidos por um indivíduo, uma função oferta de crimes:

$$o_{ij} = \delta(p_{ij}, f_{ij}, \pi_j) \quad (1.1)$$

Onde o^1 é o número de crimes do tipo i cometido pelo indivíduo² j por unidade de tempo; p é a probabilidade deste indivíduo ser condenado; f é desutilidade decorrente da punição recebida pelo crime (ou seu equivalente monetário), enquanto π representa outras variáveis que podem influir na

¹ Do inglês *offense*, mantendo a notação original de Becker

² A oferta de "mercado", dada pela soma das funções individuais, teria forma similar, tendo p , f e u substituídos pelos seus valores médios.

escolha do indivíduo entre cometer ou não um crime. Por exemplo, um aumento na renda disponível em atividades legais, ou uma melhoria em fatores como educação poderiam levar a uma redução no incentivo à atividade criminal. Ou ainda, uma modificação no tipo de punição: uma mudança de multa para prisão poderia reduzir, ainda que temporariamente, o número de crimes pois, enquanto durasse seu encarceramento não poderia cometer crimes³.

Como somente criminosos condenados são punidos, há uma “discriminação de preço” acompanhada de incerteza pois, se for condenado, ele paga f por crime em que for condenado e não paga nada se não for. A utilidade esperada do crime a ser cometido é dada por (omitindo-se, por simplicidade, os índices i e j):

$$E(U) = pU(Y-f) + (1-p)U(Y) \quad (1.2)$$

Onde Y é a renda (monetária e psicológica) obtida no crime pelo indivíduo e U a sua função utilidade.

É imediato que:

$$\frac{\partial E(U)}{\partial p} = U(Y-f) - U(Y) < 0 \quad (1.3)$$

$$\frac{\partial E(U)}{\partial f} = -pU'(Y-f) < 0 \quad (1.4)$$

Desta forma, um aumento em p ou f tenderá a reduzir o número de crimes cometidos, pois reduzirá a utilidade esperada dos mesmos, aumentando assim a probabilidade de se “pagar” um “preço” mais alto ou o próprio preço. Ou seja:

³ Note-se que não se trata aqui, de maior “dureza” ou desutilidade da pena, pois esta já está expressa na variável f , mesmo porque, em tese, é possível encontrar um valor para a multa em que o indivíduo seria indiferente a este valor e os anos a que for condenado.

$$o_p = \frac{\partial o}{\partial p} < 0 \quad (1.5)$$

$$o_f = \frac{\partial o}{\partial f} < 0 \quad (1.6)$$

Uma questão que se levanta nesta abordagem é a resposta da função o a variações em f ou p . Um aumento de p compensado por uma redução em f no mesmo percentual não reduziria a renda esperada pelo crime. Se não, vejamos:

$$E(Y) = p(Y-f) + (1-p)Y = Y - pf \quad (1.7)$$

Mas poderia ter efeito na utilidade esperada. A elasticidade da utilidade esperada em relação a p é dada por:

$$\frac{-\partial E(U)}{\partial p} \frac{p}{U} = [U(Y) - U(Y-f)] \frac{p}{U} \quad (1.8)$$

Enquanto que, em relação a f será dada por:

$$\frac{-\partial E(U)}{\partial f} \frac{f}{U} = pU'(Y-f) \frac{f}{U} \quad (1.9)$$

Colocando os dois termos lado a lado e rearranjando vem:

$$\frac{U(Y) - U(Y-f)}{f} \geq < U'(Y-f) \quad (1.10)$$

A resposta a p será maior, igual ou menor do que a f se o termo da esquerda for maior, igual ou menor do que o termo da direita. Mas o termo da esquerda é a variação média na utilidade entre $Y-f$ e Y . É fácil ver que será maior, igual ou menor do que o termo da esquerda se U'' for maior, igual ou menor do que zero. Mas a preferência ao risco é definida justamente por $U'' > 0$, neutralidade por $U'' = 0$ e aversão por $U'' < 0$.

Desta forma, se os indivíduos potencialmente criminosos apresentarem preferência ao risco, mudanças em p terão maior efeito; analogamente, mudanças em f terão maior efeito se forem aversos ao risco

e o efeito de mudanças em f ou p serão equivalentes se forem neutros em relação ao risco⁴.

A função oferta de crimes de Ehrlich(1973) difere da de Becker por incluir explicitamente o rendimento das atividades a serem escolhidas (legal e ilegal) e a possibilidade de desemprego:

$$O_{ij} = O_{ij}(P_{ij}, f_{ij}, W_{lj}, W_{lj}, U_{lj}, \pi_j) \quad (1.11)$$

Nesta equação w_i é o rendimento da atividade ilegal i , enquanto w_l é o rendimento da atividade legal e u é a probabilidade de desemprego nesta atividade. Com relação às componentes da variável π , Ehrlich inclui a riqueza pessoal ou familiar, a eficiência do indivíduo em se proteger e a quantidade de seguro privado fornecido por sua família (ou organização criminosa)⁵.

A atividade l é segura de tal modo que seus rendimentos líquidos são dados por:

$$W_l = W_l(L) \quad (1.12)$$

Onde L é o tempo gasto na atividade legal.

Já a atividade i tem risco, condicionado a dois estados da natureza. O estado "a" representa apreensão e punição com probabilidade p e o estado "b" representa sair ileso com probabilidade $1-p$. Se o estado "b" ocorrer, o rendimento da atividade ilegal, $W_i(I)$ (sendo I o tempo gasto na atividade ilegal) é integralmente recebido. Se "a" ocorre, este rendimento é descontado pelo valor da penalidade pelo crime F_i . As relações entre F e f e W e w são tais que:

$$f \equiv \frac{\partial F}{\partial t} \quad (1.13)$$

⁴ Isto é particularmente relevante na situação brasileira onde se discute a questão do endurecimento das penas como forma de desestimular a atividade criminosa. Tal medida pode ter pouco efeito dependendo do comportamento dos indivíduos em relação ao risco.

⁵ O nível do seguro é relevante por ser possível um seguro contra o desemprego.

$$w = \frac{\partial W}{\partial t} \quad (1.14)$$

A utilidade de um dado estado da natureza s é dada por:

$$U_s = U(X_s, N) \quad (1.15)$$

Onde X_s é o estoque de riqueza (que inclui o equivalente em ganhos não monetários das atividades legal e ilegal) de um dado estado da natureza s . E N é o tempo gasto em consumo ou atividades fora do mercado (lazer, por exemplo). Para cada estado da natureza, temos:

$$X_a = W_0 + W_i(I) - F_i(I) + W_l(L) \quad (1.16)$$

$$X_b = W_0 + W_i(I) + W_l(L) \quad (1.17)$$

Nas equações acima W_0 é a quantidade de riqueza já disponível para o indivíduo.

A utilidade esperada será então, dada por:

$$E[U(X_s, N)] = (1-p)U(X_b, N) + pU(X_a, N) \quad (1.18)$$

O problema consiste em maximizar a equação (1.18) sujeita às restrições de riqueza dadas pelas equações (1.16) e (1.17) e pela restrição temporal:

$$t = N + I + L \quad (1.19)$$

Além da condição de não negatividade de I e L ⁶.

É fácil ver que a alocação ótima de tempo entre as atividades deverá satisfazer:

$$(w_l - w_i)(1-p)U'(X_b) + (w_l - f_i - w_i)p.U'(X_a) = 0 \quad (1.20)$$

Este resultado pode ser generalizado para n estados da natureza. Suponhamos, por exemplo, que, além do sucesso ou fracasso ao escapar ileso da punição pelo crime (com probabilidades $1-p$ e p , respectivamente)

⁶ N é tomado como fixo.

exista ainda a incerteza de permanecer empregado na atividade legal (com probabilidade $1-u$) ou perder o emprego (com probabilidade u). Neste caso, temos 4 estados da natureza: "a" corresponde a permanecer empregado e sair impune; "b" significa sair impune, porém perder o emprego; "c" é ser punido pelo crime, mas permanecer no emprego e "d" representa a punição pelo crime acompanhada do desemprego.

Nestas condições, a utilidade esperada será:

$$E[U(X_s, N)] = (1-p)(1-u)U(X_a, N) + (1-p)uU(X_b, N) + p(1-u)U(X_c, N) + puU(X_d, N) \quad (1.21)$$

E a alocação ótima do tempo deverá satisfazer:

$$(1-p)(1-u)(w_l - w_i)U'_a + (1-p)uw_l U'_b + p(1-u)(w_l - f_i - w_l)U'_c + pu(w_l - f_i)U'_d = 0 \quad (1.22)$$

O modelo oferece uma explicação de porque pessoas que se engajam em atividades ilegais incorrem em reincidência. O criminoso provavelmente repetirá a atividade ilegal no período subsequente se as oportunidades disponíveis a ele permanecerem inalteradas. Nas palavras de Ehrlich(1973, p.529):

Reincidência não é necessariamente, portanto, o resultado de miopia, comportamento errático, ou falta de autocontrole por parte do criminoso, mas o resultado de escolha ditada por oportunidades.

Na análise de Block e Heineke(1975), há a diferença de que os autores incluem explicitamente os tempos gastos nas atividades legal e ilegal na função utilidade esperada:

$$E(U) = U(L, I, W) \quad (1.23)$$

Onde, à semelhança da escolha entre trabalho e lazer típica dos livros texto de microeconomia, a utilidade do trabalho é negativa. Assim sendo, seja este trabalho legal ou ilegal, a relação da função utilidade com o tempo gasto nestas atividades será dada por:

$$U_L \equiv \frac{\partial U}{\partial L} < 0 \quad (1.24)$$

$$U_I \equiv \frac{\partial U}{\partial I} < 0 \quad (1.25)$$

E, obviamente:

$$U_W \equiv \frac{\partial U}{\partial W} > 0 \quad (1.26)$$

Sendo r a taxa de retorno da atividade legal e V a taxa de retorno da atividade ilegal, W será dado por:

$$W = W_0 + rL + (V - pF)O(I) \quad (1.27)$$

Onde O é a quantidade total de crimes e é tal que $\frac{\partial O}{\partial I} = o$, isto é, assume-se que a quantidade total de crimes é proporcional ao tempo gasto na atividade ilegal.

A decisão do indivíduo em alocar seu tempo em atividades legais ou ilegais, isto é, a sua oferta individual de trabalho ou crime será determinada por:

$$\max_{L,I} \int U[L, I, W_0 + rL + (V - pF)O]g(p)dp \quad (1.28)$$

A função densidade de probabilidade subjetiva $g(p)$ é acrescentada por Block e Heineke indicando as crenças do agente no que se refere à chance de ser preso.

Neste caso, a condição de primeira ordem para máximo relativo é dada por:

$$E[U_I - U_L + U_W ((V-pF)O' - r)] \leq 0 \quad (1.29)$$

Onde $O' \equiv \frac{\partial O}{\partial I}$. Vamos dirigir nossa atenção a um dos termos da inequação (1.29). Não é difícil ver que, pela própria definição de covariância que:

$$E[U_w(V-pF)] = \text{cov}(U_w, V-pF) + E(U_w)E(V-pF) \quad (1.30)$$

O sinal da covariância explicitada na equação (1.30) depende de como U_w varia quando $V-pF$ varia. Se a probabilidade p de prisão aumenta (e, portanto, $V-pF$ diminui), diminui W , que por sua vez diminuirá ou aumentará U_w dependendo do sinal de U_{ww} . Se $U_{ww} < 0$ (o que significa que o indivíduo é averso ao risco) então um aumento de $V-pF$ causará uma queda em U_w , ou seja, $\text{cov}(U_w, V-pF) < 0$. Fazendo um raciocínio análogo, $U_{ww} > 0$ (preferência por risco) implica $\text{cov}(U_w, V-pF) > 0$ e $U_{ww} = 0$ (neutralidade em relação ao risco) implica $\text{cov}(U_w, V-pF) = 0$.

O comportamento em relação ao risco, portanto, não define integralmente o sinal da expressão (1.29), o que difere das abordagens anteriores de Becker e Ehrlich. Consideremos um indivíduo para o qual $U_I - U_L > 0$ (que Block e Heineke definem como um indivíduo que tem preferência pela honestidade). Se este for averso ao risco, uma condição necessária, mas não suficiente para que $I > 0$ é que os retornos da atividade ilegal sejam maiores que os custos esperados (que consistiriam na pena mais os custos de oportunidade da atividade legal não realizada durante o tempo I). Reescrevendo a inequação (1.29) é possível verificar esta afirmação:

$$U_I - U_L + [\text{cov}(U_w, V-pF) + E(U_w)E(V-pF)]O' - U_{wr} \leq 0 \quad (1.29')$$

Como, por hipótese $U_I - U_L < 0$ (preferência por honestidade) e $\text{cov}(U_w, V-pF) < 0$ (aversão ao risco), O' , U_w e r são necessariamente positivos, é necessário que $E(V-pF)$ (retornos menos custos esperados) da atividade ilegal sejam positivos. A condição suficiente para que $I > 0$ é a de que estes retornos sejam suficientemente altos para que superem a desvantagem psicológica de participar em atividades ilegais. Assim sendo, aumentos na probabilidade de prisão, da penalidade ou das oportunidades de trabalho legais irão fazer com que este grupo de indivíduos diminuam o tempo dedicado ao crime ou mesmo não cometerão crime algum.

É interessante notar, no entanto, que para indivíduos que apresentem preferência por honestidade e também por risco, tais medidas podem não afetar sua participação em atividades criminosas.

Balkin e McDonald(1981) analisam o crime através de um modelo de equilíbrio de mercado, isto é, acrescentam à análise uma "demanda", além da oferta de crimes. Para isso, partem da seguinte identidade:

$$C = Xc \quad (1.31)$$

Onde C é a taxa de crime "nominal", medida em crimes *per capita*⁷; X é o tempo de exposição ao crime escolhido pelos cidadãos (horas passadas em atividades em locais públicos⁸) e c é a taxa de crime "real". A variável c, que pela identidade acima pode ser definida como C/X, pode ser entendida como a probabilidade de se tornar uma vítima, caso exposto.

A taxa nominal de crimes pode crescer tanto porque os cidadãos aumentam sua exposição (aumentos em X) ou porque o potencial de criminalidade na comunidade aumenta (aumentos em c).

A taxa real de criminalidade também tem efeitos na exposição. Podemos escrever:

$$X = X(c, \zeta) \quad (1.32)$$

Onde o tempo de exposição será menor se o potencial criminoso aumentar, isto é:

$$\frac{\partial X}{\partial c} < 0 \quad (1.33)$$

E ζ é um vetor que contém outras variáveis que afetam a exposição. Desta forma a relação entre c e C pode ser ambígua. Diferenciando (31) em relação a c, vem:

$$\frac{\partial C}{\partial c} = X + c \frac{\partial X}{\partial c} \quad (1.34)$$

$$\frac{\partial C}{\partial c} = X \left(1 + \frac{c}{X} \frac{\partial X}{\partial c} \right) \quad (1.35)$$

⁷ A variável C é a taxa de criminalidade usual.

⁸ Segue daí, obviamente, que o estudo de Balkin e McDonald(1981) é válido, como o próprio título de seu artigo sugere, para crimes praticados "na rua", isto é, em locais públicos.

$$\frac{\partial C}{\partial c} = X (1 + \varepsilon) \quad (1.36)$$

Onde ε é elasticidade de X em relação a c. Como ε é negativo, temos duas situações possíveis: se $\varepsilon > -1$ então o aumento de c levará aumentos em C, enquanto que aumentos em c levarão à queda em C se $\varepsilon < -1$. Isto é, uma demanda "elástica" por X implica numa relação negativa entre as taxas real e nominal, o que é um resultado pouco intuitivo.

Um modelo para o comportamento da vítima pode ser explicitado pela seguinte função utilidade:

$$U = U(y, X, H) \quad (1.37)$$

Sendo y a renda líquida (isto é, descontadas as perdas pela atividade criminal) e H o tempo gasto em atividades que não levam à exposição.

O valor esperado da renda líquida é dado por:

$$E(y) = G + rL - csX \quad (1.38)$$

Em que G é a renda obtida em outras fontes que não o trabalho e s é a perda sofrida toda vez em que se é vítima de um crime.

O problema do comportamento da vítima é, do ponto de vista operacional, semelhante ao do criminoso nas formulações de Ehrlich(1973) e Block e Heineke(1975), ou seja, deve-se maximizar a função utilidade dada pela equação (37), sujeita a restrição temporal:

$$t = L + H + X \quad (1.39)$$

Considerando constantes as variáveis não relacionadas ao "mercado" de crimes, podemos sintetizar a função demanda por exposição como se segue:

$$X = X(c, s) \quad (1.40)$$

Sendo as derivadas parciais em relação a c e s ambas negativas.

Do lado da "oferta", há o criminoso, que distribui seu tempo da forma já discutida, isto é, entre atividades legais, ilegais além de tempo dedicado a atividades fora do mercado, isto é $t = N + I + L$. A diferença na abordagem de Balkin e McDonald é que o tempo gasto na atividade ilegal não é diretamente proporcional ao número de crimes, e conseqüentemente ao rendimento da atividade criminosa. Considerando a probabilidade de achar uma vítima por unidade de tempo gasto em atividades ilegais como q e assumindo que as perdas para a vítima são iguais aos ganhos para o criminoso⁹ teremos a renda esperada para este último dada por:

$$E(y) = (1-p)(G + rL + qsI) + p(G + rL + qsI - F) \quad (1.41)$$

Da mesma forma como foi feito para as vítimas, tomando como constantes as demais variáveis, a função oferta de tempo nas atividades ilegais será dada por:

$$I = I(q, s, p, F) \quad (1.42)$$

Por simplicidade, admitamos que exista apenas uma vítima e um único criminoso. O equilíbrio de mercado para a taxa nominal de crimes será dado pela igualdade da taxa nominal de crime "demandada" e a taxa nominal "ofertada", isto é:

$$cX = qI \quad (1.43)$$

Equilíbrio este composto pelas equações (1.40) e (1.42) e mais:

$$c = c(I) \quad \frac{\partial c}{\partial I} > 0 \quad (1.44)$$

$$q = q(X) \quad \frac{\partial q}{\partial X} > 0 \quad (1.45)$$

⁹ Isto pode ser válido para crimes como o roubo, mesmo assim não levando em conta a possibilidade de seguro e mesmo os ganhos ou perdas psicológicas. Certamente não é válido para crimes como o homicídio.

O equilíbrio está mostrado na figura 1.1.

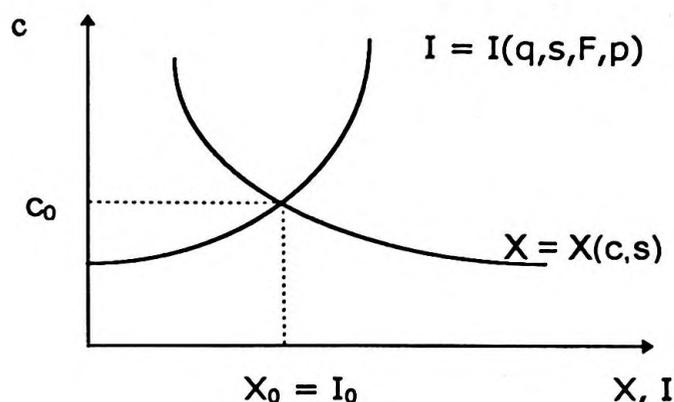


figura 1.1

O equilíbrio determina o tempo de exposição X_0 e a taxa c_0 e, conseqüentemente as variáveis q , I_0 e C .

Ehrlich(1996) expande o conceito de mercado de Balkin e McDonald(1988) à medida que a sua análise não se limita aos crimes cometidos "na rua", tendo em vista que, do lado da demanda, a questão relevante é a proteção requerida pelo indivíduo. A formulação de Balkin e McDonald pode ser considerada um caso particular em que a vítima em potencial busca proteção através de uma exposição menor.

A função de oferta de Erlich(1996) não difere significativamente¹⁰ daquela apresentada em seu artigo anterior¹¹. Podemos resumi-la, *ceteris paribus*, como uma função do retorno da atividade criminal:

$$o = o(V) \quad (1.46)$$

Onde, obviamente, $o'(V) > 0$, então temos uma oferta positivamente inclinada a exemplo de uma curva de oferta comum.

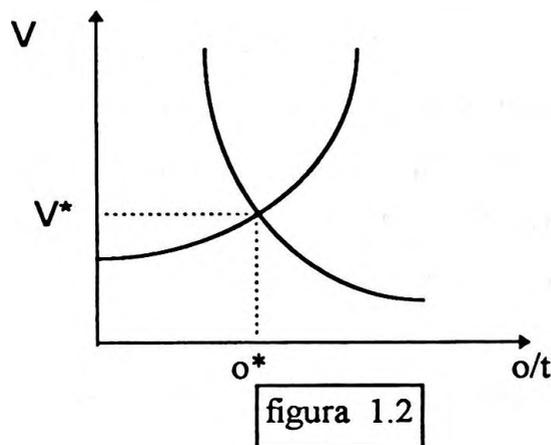
Do lado da demanda, o problema pode ser dividido em dois: a questão da demanda por proteção privada e por proteção pública.

No lado da proteção privada, esta é um custo para aquele que a contrata. Desta forma, o nível de proteção requerido será tal que o

¹⁰ A única diferença importante é que Ehrlich inclui a tendência à violência como um dos argumentos da função. É uma variação da preferência pela honestidade de Block e Heineke.

benefício marginal da proteção iguale o custo marginal do crime para a vítima. Da mesma forma, é razoável imaginar que quanto maior a taxa de crimes C , maior será a proteção requerida. A proteção maior, por outro lado, representa uma redução no retorno esperado para o criminoso, assim sendo, a relação entre a taxa de crimes e o retorno será negativamente inclinada, como numa função demanda comum, como é mostrado na figura 1.2.

Do ponto de vista público, temos que o crime é, por definição, uma externalidade e a manutenção da lei é um bem público. O nível ótimo de gastos em segurança segue o mesmo princípio de igualar benefícios marginais de redução do crime com os custos marginais de preveni-lo e reprimi-lo, com a diferença de que, no caso privado, levam-se em conta as potenciais perdas privadas com o crime enquanto que o Estado deve se ater às perdas sociais e isso inclui as interações entre mercados para diferentes crimes (por causa da substitubilidade e complementaridade existente entre diferentes tipos de crimes), bem como eventuais efeitos de propagação do crime em regiões geográficas distintas.



O mercado da figura 1.2 estará em equilíbrio para a quantidade de crime o^* que é tal que nem os criminosos, buscando um retorno líquido para o seu crime, nem os agentes privados, preocupados com o risco de se tornarem vítimas, nem o governo, tomando como base a função de bem estar social relevante, acham necessário alterar seu comportamento

Este modelo tem duas implicações importantes: uma já aparece em Ehrlich(1973) que é a questão da reincidência, que é um comportamento

¹¹ Ehrlich(1973)

esperado, ao invés de uma exceção; o outro, é que existe um nível de crime socialmente ótimo, isto é, tolerável, dados os custos de proteção e repressão. Crime é, portanto, um fato “normal” na sociedade.

Gaviria(2000) acrescenta algumas explicações para o fenômeno da criminalidade.

Uma delas é o que ele chama de “congestionamento da imposição da lei”¹², que é a relação positiva entre a probabilidade de ser punido e a taxa de criminalidade. Isto é, quanto maior o número de crimes, mais difícil se torna punir os criminosos, tomando como constantes os recursos policiais. E, como se torna mais difícil punir os criminosos (redução em p), a criminalidade aumenta. Este argumento tem implícita a suposição de que o aumento da criminalidade implica num maior número de pessoas cometendo crimes. Se o aumento de crimes se dá pelo mesmo número de pessoas cometendo mais crimes cada uma, o argumento não é necessariamente válido.

Gaviria também aponta uma inércia na taxa de crimes. Esta inércia tem diferentes causas: segundo Sah(1991) as pessoas estimam a probabilidade de ser punido baseado em uma “amostragem” de seus vizinhos, usando para isso informações passadas e presentes. Isto implica que um aumento nos gastos com segurança teriam um efeito muito pequeno no curto prazo. Uma outra fonte de inércia é apontada por Freeman, Grogger e Sonstelie(1996) que altas taxas de criminalidade hoje podem levar a jovens desistirem da educação formal em virtude de perspectivas de retornos mais altos em atividades ilegais no futuro, o que os levará a ter menos acesso ao mercado de trabalho legal.

Gaviria ainda aponta a questão da “aprendizagem” como causa da propagação do crime dentro e fora de uma região geográfica, que se dá, por exemplo, pelo convívio de criminosos em prisões. No caso específico da Colômbia, seu foco de estudo, Gaviria aponta a aprendizagem causada pelos cartéis do narcotráfico. Estes teriam um papel análogo ao de firmas multinacionais em países menos desenvolvidos, cujas firmas locais se

¹² Do original “*congestion in law enforcement*”.

beneficiam da transferência de técnicas de produção, gerenciamento e *marketing*. Da mesma forma¹³ criminosos locais poderiam se beneficiar do conhecimento dos cartéis em compra de armas em mercados negros internacionais, lavagem de dinheiro e mesmo "conexões" dentro de órgãos policiais.

Freeman, Grogger e Sonstelie(1996) apresentam um estudo de interação do crime entre áreas vizinhas. Para isso, partem de um modelo com duas vizinhanças idênticas, cada uma das quais com N pessoas idênticas. Cada pessoa decide entre ser um trabalhador e um criminoso¹⁴ e cada criminoso decide se atua (rouba) na vizinhança 1 ou 2. As pessoas, por hipótese, não mudam o seu local de residência, mas podem escolher o local de sua atividade, seja ela legal ou ilegal.

Nesta economia há apenas um bem de consumo. As pessoas são neutras em relação ao risco e obtém sua utilidade através do consumo deste bem. Cada pessoa recebe a dotação de uma unidade de tempo que pode ser usado para a produção de w unidades do bem ou para tentar um roubo. Se bem sucedido neste roubo, adquire z unidades do bem; se preso, perde todas as unidades do bem adquiridas.

A polícia investiga cada roubo, sendo a probabilidade de sucesso na apreensão do criminoso função crescente dos recursos policiais (m) e decrescente do número de roubos (n) na vizinhança em questão. Ou seja:

$$p = p(m_i, n_i) \quad (1.47)$$

Sendo $i = 1, 2$ o índice que diz a qual vizinhança nos referimos.

No que se refere ao retorno da atividade criminosa, este é função crescente do número de roubos na vizinhança, pela diminuição da probabilidade de captura. Por outro lado, ele é limitado pela riqueza disponível na vizinhança, portanto, a medida que aumenta o número de roubos deve diminuir a quantidade de bens disponíveis para novos ladrões

¹³ A analogia entre crime organizado e a Teoria da Firma é comum dentro da literatura de abordagem econômica do crime (v. Hellman e Alper(1997), capítulo 10).

¹⁴ Especificamente um ladrão na especificação do modelo de Freeman, Grogger e Sonstelie.

roubarem. O produto de um roubo z é função decrescente do número de roubos, de tal modo que:

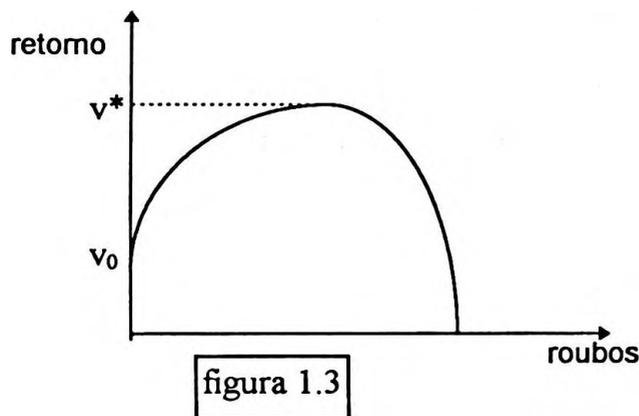
$$z = z(n_i), \quad \frac{\partial z}{\partial n_i} < 0 \quad (1.48)$$

$$\lim_{n \rightarrow 2N} z(n_i) = 0 \quad (1.49)$$

O retorno esperado (v) do crime na vizinhança i é dado por:

$$v(m_i, n_i) = (1 - p(m_i, n_i)) z(n_i) \quad (1.50)$$

O retorno é crescente quando o número de roubos é pequeno, mas após um certo limite torna-se decrescente porque o produto do roubo tende a cair, como é mostrado na figura 1.3:



Numa situação de equilíbrio interior o retorno nas três atividades deve ser igual:

$$v(n_1) = v(n_2) = w \quad (1.51)$$

E ainda que o retorno seja decrescente em ambas as vizinhanças, pois de outra forma trabalhadores e criminosos poderiam aumentar seus retornos roubando na vizinhança em que o retorno está crescendo. Há ainda a possibilidade de equilíbrios em que não haja participantes em uma ou duas atividades. Há três possibilidades

I - se $w > v^*$ então não há criminosos em nenhuma vizinhança.

II - se $v_0 < w < v^*$ pode haver um número igual de criminosos em cada vizinhança, criminosos em uma vizinhança mas não em outra ou mesmo nenhum criminoso em ambas vizinhanças.

III - se $w < v_0$ então há um número igual de criminosos nas duas vizinhanças.

1.1 O crime de homicídio

Nesta seção trataremos especificamente do crime de homicídio dentro da abordagem econômica.

Ehrlich(1973, p.532) faz o seguinte comentário com respeito aos crimes contra a pessoa, entre os quais se inclui o homicídio:

Crimes contra a pessoa são geralmente motivados por ódio ou paixão... assim, deve ser apropriado considerá-los atividades fora do mercado, isto é, atividades que visam diretamente a necessidades, em contraste com atividades de mercado ou geradoras de riqueza.

Esta afirmação é discutível, pois o homicídio pode ser motivado por razões de "mercado". Por exemplo, o assassinato de um líder de um bando rival pode levar ao aumento dos retornos de uma quadrilha dentro do crime organizado.

Mesmo numa situação em que este tipo de homicídio é uma fração insignificante, o criminoso também responderia a incentivos como a probabilidade de ser preso e o tamanho da pena. De fato, continua Ehrlich:

Já que aqueles que odeiam não respondem a incentivos de maneira diferente daqueles que amam ou são indiferentes ao bem estar de outros, a análise precedente aplicar-se-ia, com algumas modificações, aos crimes contra a pessoa assim como aos crimes que envolvem ganhos materiais.

Hellman e Alper(1997) por exemplo, separam os homicídios em "homicídios de paixão", cometidos dentro de grupos familiares, entre amantes ou resultantes de discussões entre conhecidos, crimes com forte

carga emocional cometidos “no calor do momento” dos homicídios em que há ganho monetário ou psicológico, em que a decisão de cometer ou não o crime se dá pela comparação entre os benefícios obtidos com o homicídio e os custos de praticá-lo.

No segundo caso podemos classificar as pessoas envolvidas no crime de homicídio entre “compradores” (pessoas que desejam que o crime seja cometido) e “vendedores” (pessoas dispostas a cometer o crime). Neste caso, teríamos um mercado para o homicídio semelhante a um mercado para um bem qualquer como mostra a figura 1.4:

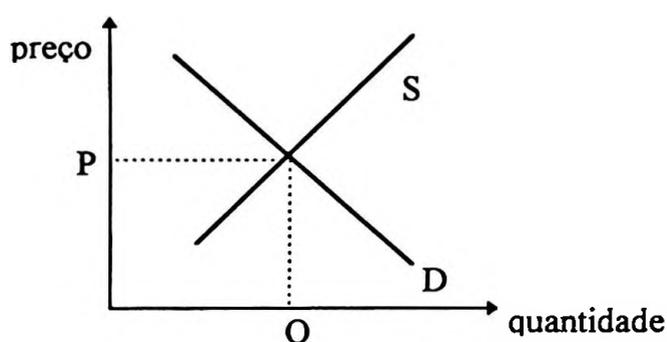


figura 1.4

A demanda é o resultado do valor monetário ou psicológico atribuído aos homicídios. À medida em que o preço aumenta, menos homicídios são “lucrativos”, isto é, menor é o número daqueles cujo ganho supera o preço, fazendo com que a demanda seja negativamente inclinada.

Já a oferta apresenta inclinação positiva, pois à medida em que os preços aumentam, aumentam o número de crimes seja porque os “produtores”¹⁵ existentes aumentam a sua “produção”, seja porque novos “produtores” entram na “indústria”.

Esta demanda pode ser deslocada por fatores similares àqueles existentes em um mercado de um bem comum: mudança nos gostos, renda, preço de bens substitutos e complementares.

Por mudança nos “gostos”, neste caso, poderíamos ter, por exemplo, o aumento do horror associado à culpa de envolvimento na morte de outra

¹⁵ O raciocínio aqui apresentado não exclui a possibilidade de que o demandante e o “produtor” sejam a mesma pessoa.

pessoa, o que levaria a um deslocamento para baixo e à esquerda da demanda.

Com relação a mudanças na renda, a análise é mais complicada, tendo em vista que, como a pena para cúmplice ou mandante do homicídio (caso do “comprador”)¹⁶ é prisão, o custo de ficar aprisionado é maior se a renda for maior. No entanto, é provável que os ganhos monetários também sejam maiores com o aumento da renda.

Como bens substitutos, extorsão e chantagem também podem ser meios de obter ganhos monetários, assim como agressão violenta pode servir para infligir ferimentos graves, tornando-se assim um bem substituto para o ganho psicológico. Quedas nos preços destes crimes podem levar a uma queda na demanda.

Do mesmo modo, é possível identificar bens complementares ao homicídio, tipicamente a informação (é preciso ter informação sobre o que está disponível, onde e a que preço) e condições de transação (tempo e risco existentes na compra do “produto”).

Com relação à oferta, assim como num mercado qualquer, será deslocada por fatores que afetem o custo que, neste caso, seriam o custo associado ao tempo gasto para execução do crime, custos materiais, psicológicos e o custo esperado da punição.

Já no caso do homicídio “por paixão” que Hellman e Alper também chamam de homicídio irracional, temos que a demanda e a oferta são totalmente inelásticas e coincidem, tendo em vista que: as pessoas que desejam que o crime seja cometido são as mesmas que o cometem, e o cometerão independente do preço, como na figura 1.5:

¹⁶ Ou do próprio homicídio, no caso de “comprador” e “vendedor” serem a mesma pessoa.

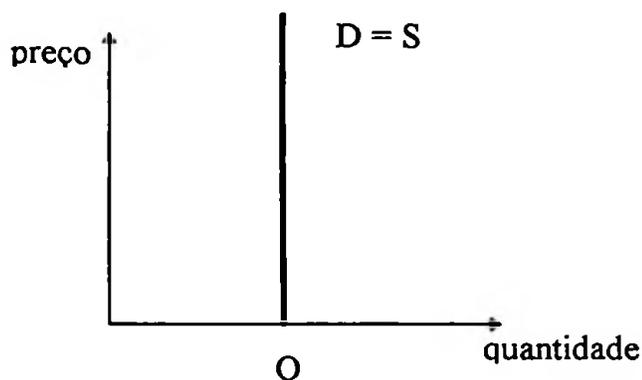


figura 1.5

Ehrlich(1975) apresenta uma abordagem mais sofisticada para a oferta de homicídios. Nesta formulação, o autor parte de duas hipóteses.

A primeira, de que este tipo de crime é cometido por motivo de ódio, ciúme e outros conflitos interpessoais envolvendo motivos pecuniários e não pecuniários, ou ainda como subproduto de crimes contra a propriedade.

A segunda é que a propensão a cometer este tipo de crime é influenciada pelas perspectivas de ganhos e perdas associadas a ele, da mesma forma que para os crimes em geral como em Becker(1968) e Ehrlich(1973).

O autor assume que a utilidade de uma pessoa i de uma perspectiva de consumo C_i depende de seu próprio consumo c_i e das atividades de consumo de outras pessoas c_j , $j = 1, 2, \dots, n$, ou:

$$U_i(C_i) = U_i(c_i, c_j) \quad (1.52)$$

Onde os sinais das derivadas parciais em relação aos c_j indicam a direção em que a utilidade de i é afetada pelas atividades de consumo de outros.

O homicídio pode ser considerado um ato deliberado do indivíduo i sobre uma vítima v de modo a tornar $c_v = 0$. O criminoso incorre em custos diretos de planejamento e execução, e enfrenta o risco de perdas em diferentes estados da natureza que envolvem aprisionamento, condenação e punição.

A condição necessária e suficiente para que o homicídio ocorra é a que a utilidade esperada de i com o crime exceda sua utilidade esperada obtida por uma ação alternativa:

$$U_i(C_i | c_v = 0) \equiv \sum_{s=a}^S p_s U_i(C_{is}) > U_i(C_i | c_v = c_{v0}) \quad (1.53)$$

Onde $s = a, \dots, S$ indica um conjunto de estados da natureza mutuamente exclusivos resultantes do homicídio.

Chamemos $p(a)$ a probabilidade¹⁷ de aprisionamento. Portanto, $1 - p(a)$ é a probabilidade de escapar da prisão.

Uma vez preso, $p(c|a)$ é a probabilidade de condenação dado o aprisionamento e $1 - p(c|a)$ é a probabilidade de ser condenado por um crime menos grave ou ser absolvido.

E ainda, se for condenado $p(e|c)$ é a probabilidade de execução¹⁸ caso seja condenado e, assim sendo, $1 - p(e|c)$ é a probabilidade de ser condenado pelo homicídio, mas não executado. Os diversos estados da natureza são resumidos no quadro a seguir:

Evento		Estado s	Probabilidade	Consumo perspectivo
aprisionamento	condenação por homicídio	execução	$p(a)p(c a)p(e c)$	$C_d: (c_i=0; c_v=0)$
		prisão por homicídio	$p(a)p(c a)(1-p(e c))$	$C_c: (c_i=c; c_v=0)$
	cond. crime menor / absolv.	outra punição	$p(a)(1-p(c a))$	$C_b: (c_i=b; c_v=0)$
não aprisionamento		sem punição	$1 - p(a)$	$C_a: (c_i=a; c_v=0)$

É evidente, pela tabela acima que $C_a > C_b > C_c > C_d$. Desta forma, estimando-se as probabilidades acima é possível investigar a eficácia das formas de punição, em particular da execução no estudo de Ehrlich(1975), sobre os homicídios.

¹⁷ Note que se tratam de probabilidades subjetivas.

¹⁸ O que é relevante para o estudo de Ehrlich, feito para os EUA, e que investiga justamente o efeito da pena capital em impedir o homicídio.

1.2 Estudos no Brasil

Com relação a estudos realizados no Brasil a respeito do crime sob a abordagem econômica podemos citar os trabalhos de Pezzin(1986), Pastore, Rocca e Pezzin(1991) e Lisboa e Andrade(2000).

No trabalho de Pezzin(1986) estimou-se equações que relacionavam os crimes (crimes contra a pessoa, contra o patrimônio e total de crimes) com variáveis sócio-econômicas como índice de pobreza, desemprego, analfabetismo, gastos com segurança pública, saldo migratório, etc. As regressões foram feitas em séries de tempo (1970 a 1984 para os valores agregados da Grande São Paulo) e *cross-section* (para municípios da grande São Paulo e subdistritos administrativos da cidade de São Paulo).

Algumas conclusões obtidas pelas estimações deste trabalho: os crimes contra a pessoa parecem ser menos influenciados por fatores sócio-econômicos do que os crimes contra o patrimônio. Isto é esperado tanto pelas formulações de Ehrlich(1975) como na de Hellman e Alper(1997), em que parte importante deste tipo de crime é do tipo "irracional" na terminologia dos últimos, motivado por "ódio e ciúme", como assinala o primeiro, ou seja, crimes com forte carga emocional.

O saldo migratório apresentou coeficiente positivo, o que corrobora a idéia de que a frustração causada pelas expectativas elevadas que tinha o migrante, associada à privação social, é um fator de incentivo à criminalidade.

O analfabetismo, no entanto, mostrou-se insignificante, o que é surpreendente pois deveria indicar menor grau de oportunidade no mercado de trabalho legal. Isto pode ser uma evidência de que, neste aspecto, prevalece a "preferência pela honestidade" de Block e Heineke(1975). Há que se observar, no entanto, que a variável não é um medidor exato do nível de educação, pois as oportunidades podem ser as mesmas para analfabetos e alfabetizados, porém com pouquíssima escolaridade, o que é um grupo mais numeroso na nossa população.

A variável gastos com segurança pública foi significativa para crimes patrimoniais, mas o mesmo não ocorreu com os crimes contra a pessoa. Nos termos de Becker(1968) e Ehrlich(1973), esta variável pode ser considerada uma *proxy* para a variável p , isto é, a probabilidade de encarceramento. Este resultado pode ter duas explicações: de novo, a possibilidade dos crimes "irracionais" terem prevalecido; segundo, maiores gastos em segurança podem ter melhorado apenas a eficiência na apreensão dos crimes patrimoniais, seja pelo maior esforço dispendido na solução destes crimes, seja pela maior dificuldade de solucionar crimes contra a pessoa.

Em Pastore, Rocca e Pezzin(1991), a investigação empírica foi bastante semelhante àquela feita em Pezzin(1986), com a diferença que no primeiro foram incluídas explicitamente as taxas de encarceramento, apresentando coeficientes negativos e significantes para todos os tipos de crime. Além disso, verificou-se uma relação negativa entre o risco de encarceramento e o número de crimes cometidos para o total de crimes e os crimes contra o patrimônio.

No que se refere às demais variáveis sócio-econômicas, os resultados foram semelhantes em ambos os trabalhos.

Já em Lisboa e Andrade(2000) é feito um estudo para homicídios nos estados de São Paulo, Rio de Janeiro e Minas Gerais, entre os anos de 1981 a 1997. O estudo é feito também para diferentes faixas etárias e é mostrado que a relação entre violência e variáveis ligadas aos ciclos econômicos (salário real e desemprego) apresenta importantes diferenças quando se considera a idade das vítimas.

Para os mais jovens, as variáveis econômicas são importantes. O salário real é significativo e negativo dos 15 aos 19 anos, assim como desemprego para a faixa dos 15 aos 20, sendo praticamente nulas para idades maiores, excetuando-se o salário real para idades acima de 30, onde o coeficiente é positivo. Ou seja, maior salário real diminui a chance de mais jovens serem vítimas de homicídio, mas aumenta-se esta chance para os mais velhos. Isto pode ser explicado pelo fato de que maiores salários reais afastam jovens de atividades criminais, onde correm maiores riscos.

O sinal do desemprego pode ser considerado surpreendente. No entanto, se considerarmos a análise de Hellman e Alper(1997), admitindo que o desemprego pode ser considerado uma *proxy* para a renda (com o sinal trocado), vemos que a relação entre crimes de homicídio e a renda é ambígua¹⁹.

No que se refere às diferenças entre os estados, os resultados evidenciam que o Rio de Janeiro é mais violento do que São Paulo, que por sua vez é mais violento do que Minas Gerais.

1.3 O modelo de propagação

Nesta seção apresentaremos o modelo cuja relação causal será testada econometricamente no capítulo 4.

A hipótese central do modelo gira em torno da probabilidade de encarceramento p , embora esta não apareça explicitamente. Será feita aqui uma hipótese semelhante à de Gaviria(2000) ou de Freeman, Grogger e Sonstelie(1996), de que p está negativamente relacionado ao número de crimes.

É preciso considerar que esta probabilidade é subjetiva, ou seja, depende da percepção dos agentes. Embora Freeman, Grogger e Sonstelie mencionem evidências empíricas, e mesmo os resultados obtidos por Pastore, Rocca e Pezzin(1991) de que a taxa de encarceramento diminui, de fato, com o aumento número de crimes, bastaria para este argumento que a percepção das pessoas fosse de que a probabilidade de ser preso diminui à medida em que o crime aumenta, o que é bastante razoável.

Consideremos, por simplicidade, apenas duas vizinhanças. Por hipótese os recursos policiais são semelhantes nas duas vizinhanças.

¹⁹ A rigor, esta ambigüidade está presente em todos os crimes, pois com o aumento da renda há aumento dos ganhos monetários do crime vis-à-vis aumento de custos de oportunidade. No caso do homicídio, entretanto, é preciso lembrar que frequentemente os ganhos associados a ele são de natureza não monetária e, por isso, não afetados pela renda.

A Utilidade esperada do indivíduo que reside na vizinhança 1, mas também pode cometer um crime na vizinhança 2 é dada por:

$$E(U_t) = U(O_{1,t}, O_{2,t}, p_{1,t}, p_{2,t}) \quad (1.54)$$

Considerando o rendimento das atividades legal (W), criminal (V) e a punição (F) como constantes. A função expressa na equação (1.54) pode ser reescrita na forma:

$$E(U_t) = W + VO_{1,t} - p_{1,t}F + VO_{2,t} - p_{2,t}F \quad (1.55)$$

O indivíduo irá maximizar esta utilidade sujeito à restrição de tempo:

$$t = N + L + I_1 + I_2 \quad (1.56)$$

Isto é, o tempo será dividido entre atividades de consumo, atividades de trabalho legais e atividades de trabalho ilegais nas duas vizinhanças.

Se, na vizinhança 1 no período $t-1$ ocorre um aumento da criminalidade, *ceteris paribus*, a criminalidade no período t será ainda maior porque o aumento em $t-1$ diminui a probabilidade de encarceramento p . Ou seja:

$$O_{1,t} = O_{1,t}(p_{1,t}), \quad \frac{\partial O_{1,t}}{\partial p_{1,t}} < 0 \quad (1.57)$$

$$p_{1,t} = p_{1,t}(O_{1,t-1}), \quad \frac{\partial p_{1,t}}{\partial O_{1,t-1}} < 0 \quad (1.58)$$

O que implica numa relação positiva entre a criminalidade no período t com ela mesma no período $t-1$.

Resta saber a relação entre vizinhanças. É razoável supor que o aumento na criminalidade da vizinhança 1 no período $t-1$ também irá afetar a criminalidade da vizinhança 2 no período t ? A resposta é sim, pois imaginemos que as vizinhanças são suficientemente pequenas de tal modo que o criminoso ao expandir sua atividade²⁰ "invada" a vizinhança 2, seja

²⁰ O raciocínio pode ser expandido para indivíduos residentes na vizinhança 2 se considerarmos a hipótese de "aprendizagem" de Gaviria(2000).

porque esta é uma conseqüência natural desta expansão, já que as vizinhanças são pequenas, seja porque ele percebe que os recursos policiais não diferem significativamente de uma vizinhança para outra e que, portanto, a polícia terá aumentadas suas dificuldades com o aumento do crime (não necessariamente na mesma magnitude). De tal modo que podemos escrever:

$$p_{i,t} = p_{i,t}(O_{2,t-1}, O_{1,t-1}) \quad i=1,2 \quad (1.59)$$

Com as duas derivadas parciais sendo positivas. Ou seja, mantidos os recursos policiais e outras variáveis que determinem a probabilidade de apreensão constantes, um aumento da criminalidade numa vizinhança afetará a probabilidade de encarceramento em **ambas** as vizinhanças.

Como a criminalidade em uma vizinhança é função decrescente da probabilidade de encarceramento naquela vizinhança:

$$O_{i,t} = O_{i,t}(p_{2,t}), \quad \frac{\partial O_{i,t}}{\partial p_{i,t}} < 0 \quad (1.60)$$

Combinando as equações (59) e (60) para $i=2$, temos:

$$O_{2,t} = O_{2,t}(O_{2,t-1}, O_{1,t-1}) \quad \frac{\partial O_{2,t}}{\partial O_{1,t-1}} > 0 \quad (1.61)$$

E, de maneira análoga para $O_{2,t}$.

Estabelece-se aí, uma relação causal, tanto no espaço como no tempo. Ou seja, o aumento na criminalidade na vizinhança 1 no período $t-1$ causando o aumento da criminalidade na vizinhança 2 no período t . Este modelo pode ser generalizado para n regiões vizinhas e uma defasagem de P períodos de tempo.

$$O_{i,t} = O_{i,t}(O_{j,t-p}), \quad p = 1,2,\dots, P, \quad j = 1, 2,\dots, n \quad (1.62)$$

2. Causalidade

Hicks (1979) dizia que: "causalidade e economia... são palavras que não são freqüentemente encontradas juntas". Segundo ele, os economistas, embora preocupados com efeitos e, talvez menos freqüentemente, com causas, deixariam a relação entre causa e efeito a cargo dos filósofos.

Isto não chega a ser surpreendente. Um dos motivos é a própria natureza da teoria econômica, onde existem poucas leis para as quais há aceitação por parte de todos economistas. Além disso, o conhecimento econômico é, pela sua própria natureza, "imperfeito", já que poucos fatos econômicos são conhecidos precisamente, sendo suas medições sujeitas inexoravelmente a erros e, em muitos casos, a ambigüidades, como, por exemplo muitas das variáveis macroeconômicas (emprego, produto, etc.).

Há um outro motivo, que é a relação entre a ciência econômica e o tempo. Embora, mesmo nas ciências naturais, o conhecimento se altere ao longo do tempo²¹, esta é uma transformação que decorre, exclusivamente, do próprio conhecimento adquirido. Um experimento em Física ou em Química produziria resultados iguais em qualquer época, se fosse conduzido da mesma forma. Embora isto não seja uma verdade absoluta, é certamente correto para o período de tempo em que a espécie humana existe.

Isto, entretanto, não é válido para a Economia, ou tampouco para qualquer ciência social, tendo em vista que estas de fato mudam com o tempo. A Economia do período feudal, do início do século e a atual, são, para citar um exemplo, três objetos de estudo distintos. As diferentes visões que teriam estudiosos colocados nestes três períodos seriam, apenas em parte, decorrente apenas da diferença do volume de conhecimento adquirido em cada época. Um físico quântico, entretanto, se colocado no

²¹ E esta alteração pode ser uma "mudança de paradigma" no sentido de Kuhn(1995), na qual poderíamos citar as "três físicas": a Física Aristotélica (não inercial); a chamada Física Clássica, iniciada em Galileu e Copérnico e que se fundamenta nas três leis de Newton e nas quatro leis de Maxwell e a Física Moderna, inaugurada pela mecânica quântica de Planck e a teoria da relatividade de Einstein, no início do século XX. Pode ser também de caráter tecnológico, como a biologia molecular após a invenção do microscópio eletrônico.

século passado, não teria nenhuma razão para desconsiderar seus conhecimentos²².

Embora a afirmação de Hume(1777) de que “é essencial a causas e efeitos... a antecedência no tempo da causa em relação ao efeito” seja um tanto discutível, pois pode haver relações onde a causa e efeito sejam simultâneas, é inegável que esta relação tenha alguma referência ao tempo. E, o fato de os economistas estarem primordialmente preocupados com os problemas do seu tempo, isto é uma explicação, mas não uma justificativa da pouca existência da questão da causalidade na literatura econômica apontada por Hicks. Isto porque, embora os problemas enfrentados pelos economistas não tenham relação com o passado remoto²³, ou terão conseqüências para o futuro longínquo, terão relação com o passado próximo e com o futuro próximo. E, próximos ou não, continuam sendo passado e futuro, portanto a relação de causa e efeito é, sim, relevante para o economista.

2.1 Definições de causalidade

A definição mais comum de causalidade é a de uma relação assimétrica entre duas variáveis, onde uma delas (causa), gera a segunda (efeito).

Esta definição traz alguns inconvenientes, como o apontado por Hume (1777), decorrente do problema de se postular a existência de uma relação entre estas duas variáveis, em outras palavras, qual seria a base lógica que permite postular tal relação, já que somente a observação do comportamento das mesmas não permite inferir a existência desta relação.

²² Há uma série de exemplos nas ciências naturais de pesquisadores que estavam, por assim dizer, “fora de seu tempo”. Um grande exemplo disso é o monge agostiniano Johann Gregor Mendell, que formulou as três leis básicas da genética em meados do século XIX, tendo seu trabalho inicialmente rejeitado e esquecido, só resgatado no começo do século XX, quando três cientistas chegaram a conclusões idênticas. Não faria sentido, no entanto, um contemporâneo de Adam Smith prever as crises de 1929 ou as crises cambiais do final do século XX.

²³ Pelo menos não relevante para eles, economistas. O mesmo não se aplica, por exemplo, aos historiadores.

Em outras palavras, a relação de causa e efeito existente entre duas variáveis decorre muito mais da percepção do observador do que de um arcabouço lógico. Não é possível afirmar, *a priori*, numa relação estabelecida entre duas variáveis X e Y, de que X seja, de fato, uma causa de Y. Embora a conexão entre ambas possa estar firmemente estabelecida, elas podem, na verdade, ter uma causa comum²⁴.

Para Hume, portanto, não é possível diferenciar uma relação causal de uma relação qualquer existente entre duas variáveis. Não seria possível, desta forma, distinguir uma previsão feita por um ou outro tipo de relação.

Para Simon(1970), este é um problema bastante sensível no que concerne os economistas, pois para estes é bastante diferente a predição num e noutro caso. O postulado de que "o vôo dos pássaros precede a chuva" é substancialmente diferente do postulado de que, "*ceteris paribus*, o aumento do preço de um bem reduz a quantidade demandada do mesmo". Em outras palavras, ao economista é necessária a distinção entre relações quaisquer entre duas variáveis e relações que sejam verdadeiramente de causa e efeito.

Um outro problema decorrente da definição de causalidade é justamente como se dá esta percepção pelos agentes. A atribuição de uma relação causal não é, necessariamente, consequência de inferência ou indução de uma seqüência de atos observados independentemente. Por exemplo, o fato de uma lâmpada acender logo após uma outra pode gerar a percepção que o acendimento da primeira causou o da segunda.

Sobre esta questão, o veredicto, não muito animador, é dado por Granger(1988), de que, de fato, pouco se sabe a respeito de como os agentes percebem a existência de uma associação ou conexão entre as variáveis.

²⁴ Este tipo de ambigüidade aparece com freqüência nas ciências sociais. Há um exemplo clássico na história brasileira, que é a participação dos bancos ingleses no financiamento da Guerra do Paraguai. Para alguns historiadores, isto é uma evidência da participação direta da Inglaterra no conflito. Para outros, no entanto, não havia este interesse, mas apenas o interesse comercial dos bancos diante de uma oportunidade de ganho. Para os primeiros, existe uma relação de causa e efeito, enquanto para os últimos são apenas dois fatos que tiveram uma causa comum (V. Doratioto, 1991).

Há ainda um terceiro problema com a definição de causalidade relacionado com a assimetria existente em uma relação causal: se X causa Y, isto não implica que Y cause X, ou mesmo que a ausência de Y implique na ausência de X. Há que se procurar, portanto, formas de garantir a existência desta assimetria em uma relação entre estas variáveis para que esta possa ser caracterizada como causal. Esta assimetria, segundo Granger, pode ser garantida em uma ou mais das seguintes formas abaixo:

- controlabilidade;
- existência de uma teoria relevante;
- conhecimento *a priori*;
- antecedência temporal.

Para muitos economistas é a controlabilidade o fator essencial para se caracterizar uma relação entre X e Y como sendo causal. Ou seja, se é possível controlar Y controlando X e não for necessariamente possível controlar X através de Y, conclui-se que X é causa de Y. O apelo desta forma para o economista é notório se imaginarmos sua relevância do ponto de vista dos formuladores de política econômica. Para muitos, no entanto, o conceito de controlabilidade seria mais forte do que o de causalidade, não havendo, portanto, equivalência, entre estes conceitos.

A questão da teoria é delicada especialmente no que diz respeito à Economia, como já foi mencionado no início do capítulo, em função da não existência de consenso em muitas das teorias econômicas. Este é um problema menor nas ciências naturais (há pouca dúvida, por exemplo, na relação de causa e efeito entre a o campo gravitacional da Lua e as marés²⁵). Mesmo quando a teoria, em si, não é objeto de grandes discussões, a questão de causa e efeito nem sempre é muito clara, haja vista a existência abundante na teoria econômica dos chamados modelos

²⁵ Embora há que se notar que, mesmo nas ciências naturais este não é um problema inexistente. Veja por exemplo, as concepções de Lamarck e Darwin de como se deu a evolução das espécies. Em ambos, na verdade, prevalece o conceito da sobrevivência do mais apto, mas o mecanismo que causa essa maior aptidão se desenvolve é que varia. Para Lamarck, a própria necessidade é a causa, enquanto que, em Darwin, este aparecimento é aleatório.

simultâneos, do tipo oferta e demanda ou IS-LM. Preço causa quantidade ou vice-versa?

Esta questão poderia ser respondida havendo o tal "conhecimento *a priori*". Se é um mercado, por exemplo, onde há restrição às importações por quotas, é bem provável que quantidade cause preço. Por outro lado, num mercado onde os preços são controlados, deve ocorrer o inverso. De um modo geral, no entanto, ambos "se causam" mutuamente e, portanto, esta é uma ferramenta limitada do ponto de vista do economista.

Finalmente, sobra a questão da antecedência temporal. Como veremos na seção 2.2, não há uma relação biunívoca entre antecedência temporal e causalidade. Entretanto, esta é uma forma de garantir a assimetria²⁶ bastante útil ao economista e em particular ao econometrista.

Em particular, os testes de causalidade mais usados na Econometria se baseiam nesta assimetria, em particular o teste de Granger, em que, aliás, se baseia este trabalho. Granger baseia o seu conceito de causalidade em dois axiomas

1º) a causa precede o efeito²⁷.

2º) a causa contém alguma informação a respeito do efeito.

Ou seja, se A_t é o conjunto de informações disponíveis no tempo t , $A_t - X_t$ este mesmo conjunto de informações exceto aquelas contidas em X_t , C é uma região qualquer e se:

$$P(Y_{t+1} \text{ em } C \mid A_t) \neq P(Y_{t+1} \text{ em } C \mid A_t - X_t) \quad (2.1.1)$$

Então dizemos que X causa Y . Em outras palavras, X causa Y se Y pode ser previsto de maneira mais eficiente utilizando as informações contidas em X_t .

²⁶ Existiu na década de 70 um debate entre Friedman e Tobin a respeito de controlabilidade e antecedência temporal no caso da oferta monetária. (V. Freitas, 1993)

²⁷ À maneira de Hume. Ou seja, a assimetria se dá pela antecedência temporal

2.2 Tipos de causalidade

Podemos definir dois tipos de causalidade. Considerando A um evento que ocorra num determinado instante²⁸ do tempo t_A e B um evento no instante t_B , sendo $t_B > t_A$, então dizemos que, se A causa B, mas é uma das possíveis causas de B, podendo haver outras, há causalidade fraca; entretanto, se A causa B, e a única razão de B ter ocorrido é a ocorrência prévia de A, então dizemos que há causalidade forte.

Entre os conjuntos de eventos definidos como de causalidade fraca há ainda dois subconjuntos: causas separáveis e não separáveis. Se A é uma causa separável de B, é concebível uma situação em que A não ocorre, mas todos os outros eventos (todas as outras causas potenciais de B) ocorrem da mesma forma no instante t_A .

Se considerarmos duas causas de B, denominadas A_1 e A_2 , separáveis, a situação hipotética mencionada anteriormente, se apresenta de três formas possíveis: se A_1 não ocorreu, mas todo o resto permaneceu inalterado, que denominaremos não- A_1 ²⁹; se quem não ocorreu foi A_2 , então não- A_2 ; e, finalmente, uma terceira possibilidade de que nem A_1 , nem A_2 tenham ocorrido, que será denominada não- A_{12} .

O quadro que se segue ilustra as possibilidades em que há, em princípio, duas causas separáveis. Em cada um dos casos, as hipóteses não- A_1 , não- A_2 e não- A_{12} são apresentadas, e onde há "sim" significa "B ocorreu" e, evidentemente o "não" significa "B não ocorreu".

²⁸ Embora tenhamos utilizado o termo "instante", este pode ser perfeitamente substituído pelo termo "período", dependendo do caso. Por exemplo, a Guerra do Paraguai pode ser considerada uma das causas do endividamento externo brasileiro no século passado e foi, portanto, uma causa que não ocorreu em um instante, mas prolongou-se por um período de alguns anos. (V. Doratioto, 1991)

²⁹ Ou, em linguagem de Teoria dos Conjuntos, o complementar de A_1 .

caso	não- A_1	não- A_2	não- A_{12}
1	não	não	não
2	não	sim	não
3	sim	não	não
4	sim	sim	sim
5	sim	sim	não
6	não	sim	sim
7	sim	não	sim
8	não	não	sim

No caso 1, B não ocorre a não ser que ambos, A_1 e A_2 ocorram. As causas são, nesta situação, aditivas. No caso 2, A_2 é irrelevante; para B ocorrer, basta que A_1 ocorra. O caso 3 é o inverso, a causa relevante é A_2 . E, no caso 4, nenhuma das duas supostas causas o é, de fato. Cumpre notar que, estes três últimos são casos particulares de causas aditivas, em que uma delas (ou ambas) tem efeito nulo.

O caso 5 pode ser chamado de "causas sobrepostas": B ocorre se uma ou outra causa ocorrer³⁰. Nos casos 6, 7 e 8 B ocorre apesar do "sim" na coluna do não- A_{12} , ou seja, B ocorre mesmo se as duas causas mencionadas não acontecerem. Portanto, é absolutamente necessário que exista uma terceira causa A_3 , que guarda, com relação às duas primeiras, relação de sobreposição. Aqui, há uma analogia com a função de produção: se há apenas dois fatores, como na clássica função que tem como argumentos capital e trabalho, eles devem ser complementares. Se há três, entretanto, pelo menos um par deve ser de fatores substitutos.

Passemos então, à questão das causas não separáveis. Suponhamos, desta vez, que A e A^* sejam causas de B, e que uma não possa ocorrer sem que tenha ocorrido a outra, caracterizando-as, assim, como causas não separáveis.

Não faria sentido a repetição da tabela acima, pois neste caso, não-A implica não-A*. Da mesma forma, não-A* implica não-A. Não é possível considerar, portanto, alternativas em que só uma delas esteja presente, como fizemos anteriormente. Só há duas alternativas possíveis, ou ambas estão presentes, ou ambas estão ausentes.

A questão que surge neste momento é: a relação entre A e A* não se caracteriza, ela mesma, por ser uma relação de causalidade, tendo em vista que a negação de uma implica a negação da outra? Se tomarmos o conceito de Hume(1777), mencionado anteriormente, de que a causa precede o efeito, a resposta teria de ser não. Contudo, como foi também mencionado, esta afirmação é discutível, e, de fato, Kant(1781) considera que a relação entre A e A* guarda uma analogia com a relação de causalidade, que ele chamou de reciprocidade, ou, numa tradução mais literal, "causalidade mútua"³¹. Esta "causalidade mútua" seria, portanto, um tipo de causalidade e, se relaxamos a afirmação de Hume de "causa *tem* que preceder o efeito" para "efeito *não pode* anteceder a causa"³², teríamos este tipo de causalidade caracterizado por ser contemporânea (causa e efeito são simultâneos).

³⁰ Seria o caso do exemplo dado por Hicks, de uma vítima que morre após ingerir dois venenos, cada um dos quais letais e que, portanto, isoladamente produziriam o mesmo efeito

³¹ O termo em alemão utilizado por Kant no original é *Weschselwirkung*.

³² Há um conhecido paradoxo na Física, muito utilizado por escritores de ficção científica, a respeito de viagens no tempo, conhecido como o paradoxo do avô: se uma pessoa voltasse no tempo e assassinasse seu avô, evitaria seu próprio nascimento, assim sendo, não voltaria no tempo e nasceria normalmente. Este tipo de paradoxo evita o que poderíamos chamar de causalidade inversa, ou seja, o efeito precedendo a causa. Alguém poderia argumentar que este tipo de causalidade ocorre na Economia, como consequência das expectativas, como por exemplo, a expectativa de que um banco não terá liquidez suficiente leva a saques volumosos dos clientes deste banco, fazendo com que o banco não tenha liquidez, o que os economistas gostam de chamar de "previsão auto-realizável". Há que se distinguir, no entanto, que as causas aqui são as expectativas e não o fato em si, portanto não se trata de efeito precedendo a causa.

3. Espaço e Tempo

Os conceitos de espaço e tempo variam em cada período histórico, ou mesmo em cada campo de atividade. A interação entre os dois conceitos também apresenta diferentes aspectos em cada ramo do conhecimento. Esta interação é por vezes sutil, mas em outras chega até a estar relacionada, ou mesmo influir toda uma estrutura de conhecimento.

Por exemplo, causa estranheza em alguns o fato de que as evidências dos primeiros povoamentos humanos do continente americano tenham sido encontradas justamente em regiões das mais áridas do nordeste Brasil. Esta estranheza, na verdade, não se justifica, pois esta região na época era justamente o ponto de contato entre a Mata Atlântica e a Floresta Amazônica. O equívoco aí, está em associar uma região do espaço a um ponto fixo no tempo.

O exemplo acima está entre aqueles casos "sutis" de interação entre espaço e tempo. Na física moderna encontramos uma relação mais radical, onde a Teoria da Relatividade de Einstein derrubou o conceito de espaço e tempo independentes.

Na Economia não é diferente, a questão do espaço e do tempo está presente em vários de seus campos. Até na Microeconomia mais básica, deve-se notar que um bem não está completamente definido sem que sejam definidas suas dimensões espaciais e temporais. Isto é, um mesmo bem, considerado apenas em suas características intrínsecas, é diferente dependendo de onde e quando está. A Microeconomia, a rigor, não distingue os "fabricantes" dos "comerciantes", sendo todos considerados produtores. Qual é, então, a participação dos comerciantes no processo de produção? É justamente modificar o tempo e o espaço em que o bem se encontra.

Apesar disso. Há que se notar que, ao longo do pensamento econômico, maior ênfase tenha sido dada ao tempo em detrimento do

espaço³³. Esta ênfase também se reflete na Econometria, que tem na análise de séries de tempo um de seus campos mais férteis, muito maior quando comparado com a Econometria Espacial.

Ao longo deste capítulo, discutiremos alguns conceitos de espaço e tempo, na Economia e em outras ciências.

3.1 Conceito de tempo

O conceito de tempo está entre aqueles denominados de conceitos primitivos. Não há, a rigor, uma definição de tempo formalmente explicitada. Isso não impede, porém, a existência de controvérsias em relação ao mesmo.

Intuitivamente, para qualquer pessoa só existe “um único” tempo, que como afirmou Newton em 1687: “absoluto, verdadeiro e matemático tempo, de si mesmo, e de sua própria natureza, flui igualmente sem relação com nada externo a ele”. O conceito newtoniano de tempo é de um tempo totalmente independente, absoluto como ele mesmo afirmou.

Kant (1781) contestou o conceito de Newton de tempo absoluto, objetivo, pois segundo ele o tempo seria uma forma ou fundamento subjetivo de todas as experiências. Ainda que subjetivo, o tempo de Kant não diferia do de Newton no sentido de ser universal, isto é, o mesmo para todos.

De fato, este caráter homogêneo do tempo não foi questionado até o final do século XIX. Uma inconsistência interessante nas leis da Física acabou levando a reformulação desta idéia, através da Teoria da Relatividade de Einstein.

O “tempo da Física”, entretanto que mais influenciou o pensamento econômico não é esse da relatividade de Einstein, nem tampouco o da Mecânica ou o do Eletromagnetismo, mas o da Termodinâmica.

³³ A principal exceção são justamente áreas ligadas a Microeconomia, especialmente em modelos de localização no espaço como o de Hotelling(1929).

A Termodinâmica, através das suas leis, mostra como certas situações são unidirecionais. Em particular a lei da Entropia, que basicamente afirma que os sistemas tendem a "se desorganizar", o que, num sentido mais restrito se refere ao equilíbrio térmico. Se colocamos um corpo quente junto a um corpo frio, eles tenderão a trocar calor e entrar em equilíbrio térmico. As moléculas mais agitadas do corpo quente chocar-se-ão com as do corpo frio até que atinjam estado de agitação semelhante. Há uma clara analogia com o movimento de capitais indo para onde houver maior retorno. Se houver um ativo A com maior retorno do que um ativo B, ambos com o mesmo nível de risco, capitais sairão de B para A equalizando os retornos.

Esta direção única dada pela Termodinâmica tem relação estreita com o tempo, através do conceito de "flecha do tempo", que só percorre uma direção. Assim, da mesma forma que um suco gelado não fica mais frio se colocado ao sol, lagos não congelam no verão, este conceito pode ser estendido a uma xícara que se quebra, mas não se junta e a própria expansão do universo. Todos estes eventos teriam relação direta com a "flecha do tempo", em virtude de sua irreversibilidade.

Hsieh e Ye(1991) chegam a afirmar que Marshall tinha conhecimento da Termodinâmica, que poderia ser identificada pela passagem:

O homem não pode criar coisas. No mundo mental e moral, ele pode sim produzir novas idéias; mas quando se diz que produz coisas materiais, ele está apenas produzindo utilidades; ou, em outras palavras, seu esforços e sacrifícios resultam em mudar a forma ou arranjo da matéria para adaptá-la melhor à satisfação dos desejos.

Embora Marshall não faça referências explícitas à Termodinâmica, não é incomum encontrar autores que o fazem, como Georgescu-Roegen(1971) e Wilson(1970). Enquanto o último utiliza o conceito de entropia para aplicações em modelos de Economia Regional e Urbana, o

primeiro faz uma crítica à teoria microeconômica tradicional e, entre outras coisas, incorpora explicitamente o tempo³⁴ em sua análise.

3.2 Conceito de espaço

O campo da matemática mais ligado às questões do espaço é, obviamente, a Geometria que, na sua forma mais tradicional, se fundamenta nos postulados de Euclides. A partir destes postulados que, à primeira vista, parecem ser indiscutíveis, são derivados os teoremas que compõem o corpo da Geometria Euclidiana, ela que lidava com um espaço de duas ou três dimensões, e que por mais de dois mil anos foi considerada a única verdadeira geometria do espaço real.

Ao longo do século XIX, no entanto, os princípios euclidianos começaram a ser contestados, em particular seu quinto postulado que dizia o seguinte: dado um plano que contém uma reta e um ponto (fora dela), por este ponto, é possível construir apenas uma reta paralela. Vários geométricos não euclidianos propuseram postulados alternativos e, a partir deles, construíram "novas" geometrias.

Em 1830, o matemático russo Nikolay Lobatchevsky anunciou uma geometria bidimensional em que um número infinito de retas poderia ser desenhado em qualquer ponto paralelas a uma outra reta do mesmo plano. Na sua geometria, a soma dos ângulos internos de um triângulo é menos do que 180° .

Em 1854, o alemão, também matemático, Bernhard Riemann desenvolveu outro espaço bidimensional em que a soma dos triângulos era maior do que 180° .

O espaço de Riemann era elíptico, enquanto o espaço de Lobatchevsky era hiperbólico, em contraste com espaço plano da geometria euclidiana. Ao final do século XIX, diferentes espaços tinham sido

³⁴ A função de produção de Georgescu-Roegen incorpora o tempo em seus argumentos.

concebidos com os mais diferentes formatos (rosquinhas, interior de um túnel, etc.

Paralelamente, enquanto prevaleciam os postulados de Euclides na Geometria, na Física o conceito de espaço seguia os mesmos caminhos.

O próprio Keynes, em sua Teoria Geral, mostra familiaridade com a questão dos espaços curvos, e possivelmente da própria Teoria da Relatividade³⁵, como mostra o trecho abaixo:

Os teóricos clássicos parecem geômetras euclidianos em um mundo não euclidiano que, descobrindo pela experiência que retas aparentemente paralelas freqüentemente se encontram, as censuram por não se manterem separadas como único remédio para as infelizes colisões que ocorrem. Na verdade, não há remédio a não ser jogar fora o axioma das paralelas e desenvolver uma geometria não euclidiana. Algo similar é preciso na Economia. Devemos jogar fora o segundo postulado da doutrina clássica e desenvolver o comportamento de um sistema em que o desemprego involuntário no senso estrito é possível.

Hsieh e Ye(1991) afirmam que, dada a citação acima, "não seria um exagero afirmar que a teoria de desemprego involuntário de Keynes foi inspirada por Einstein".

A concepção de espaço também difere em outros campos do conhecimento³⁶. Na Biologia, explora-se a diferente perspectiva de espaço em animais, enquanto na sociologia se estudam as organizações espaciais em diferentes culturas.

3.3 Espaço na economia

O pensamento econômico, pelo menos no seu eixo mais tradicional, deu mais importância ao tempo do que ao espaço. Isso se deve ao que

³⁵ V. Hsieh e Ye(1991).

³⁶ V. Kern(1983).

Isard(1956) chama de "viés anglo-saxônico", que pode ser ilustrado por este trecho de Marshall em seu *Principle of Economics*:

As dificuldades do problema residem principalmente nas variações na área do espaço, e no período do tempo pelo qual o mercado se estende, sendo o tempo mais fundamental do que o espaço.

E este "preconceito" se propagou pela economia e relegou o espaço a um segundo plano, enquanto que modelos dinâmicos surgiram em abundância na literatura econômica. Funções de demanda, oferta, propensões a consumir, aparecem como grandezas não localizadas. Segundo Isard(1956), a exceção é a chamada escola de pensamento da concorrência monopolística, particularmente Chamberlain(1933). Ponsard(1954), no entanto, tem uma visão diferente, pois considera que mesmo a teoria da concorrência monopolística reduz o espaço como elemento da categoria das causas da diferenciação do produto.

Para Isard, a primeira tentativa de construir uma teoria geral de locação é devida a Alfred Weber³⁷. Seu trabalho busca explicar as forças que operam quando pessoas ocupam uma região não desenvolvida e estabelecem um sistema econômico isolado.

Primeiramente, um estrato agrícola se forma para produzir os meios de subsistência. A área com a população agrícola serve de base geográfica para os outros estratos: o segundo é uma indústria primária que produzirá para o estrato agrícola. O terceiro estrato consistirá de indústrias secundárias. Finalmente, o quarto estrato será constituído de pessoas com funções de gerenciamento e organização geral.

Entre aqueles autores pertencentes à chamada escola da concorrência monopolística, podemos citar, além de Chamberlain Hotelling, Enke, Lösch, Hoover³⁸.

³⁷ *Über den Standort der Industrien*, Tübingen, 1909.

³⁸ V.Ponsard(1954)

3.4 Econometria Espacial

O termo foi inicialmente introduzido por Jean Paelinck no início dos anos 70 para denominar a área do conhecimento que lida com a estimação e teste do modelos econométricos multi - regionais³⁹.

A existência de uma área da Econometria denominada de Econometria Espacial⁴⁰ se justifica por dois aspectos: o primeiro é o fato da importância da questão espacial inerente à ciência regional e, em particular, à economia regional.

O segundo é que, dados distribuídos no espaço podem apresentar dependência ou heterogeneidade em sua estrutura. No que se refere à metodologia econométrica tradicional, a presença destes efeitos pode tanto requerer alguma modificação na mesma como mesmo invalidá-la. Em alguns casos é necessário mesmo a criação de novas técnicas para o correto tratamento destes efeitos. Como nota Anselin (1988), "estas questões são tipicamente ignoradas pela econometria tradicional e compõem o campo separado da Econometria Espacial".

Paelinck e Klaasen (1979) apontam as principais características do campo da Econometria Espacial no que se refere ao tipo de questões a serem consideradas:

- o papel da interdependência em modelos espaciais.
- a assimetria nas relações espaciais.
- a importância de fatores explicativos localizados em outros espaços.
- diferenciação entre interação *ex post* e *ex ante*.
- modelagem espacial explícita.

³⁹ Anselin (1988).

⁴⁰ Há quem prefira o termo "Econometria Regional", para evitar confusões com o termo espacial. Esta denominação é adequada se levarmos em conta que as principais aplicações se dão no âmbito da economia regional e urbana. Há que se notar, no entanto que seus métodos podem, em princípio, ser aplicados a qualquer tipo de espaço, não necessariamente um espaço geográfico, como por exemplo um ativo e sua localização no espaço dos retornos médios e da volatilidade.

Assim como existe a econometria espacial, existe também o campo da estatística espacial. Anselin (1988) nota que a distinção entre estes dois campos é sutil, sendo os métodos de uma utilizados por outra. Segundo ele, mais prático seria deixar a cargo dos próprios autores referirem seus trabalhos a um ou outro campo. Assim sendo, a econometria espacial tenderia a partir de um modelo ou teoria em particular e teria como foco, principalmente, a economia regional e urbana, enquanto a estatística espacial trataria primordialmente de fenômenos naturais, ligados principalmente a campos como a biologia e geologia.

Créssie(1993), no entanto, divide a estatística espacial segundo o modo em que os dados são apresentados. A econometria espacial, em particular, está mais intimamente associada, a uma destas divisões, que serão discutidas na seção seguinte.

3.4.1 Estatística Espacial

Segundo Cressie(1993), as primeiras manifestações de estatística para dados espaciais teriam aparecido na forma de mapas de dados. Há mapas deste tipo que remontam ao século XVII⁴¹.

Só bem mais tarde teriam aparecido os modelos espaciais. Um exemplo encontramos em Student (1907) que, interessado na distribuição de partículas por um líquido, descobriu que a distribuição do número de células por unidade de área seguia uma distribuição de Poisson.

Um modelo espacial pode abranger uma classe de problemas bastante grande: os dados podem ser contínuos ou discretos, podem ser dados agregados espacialmente ou observações em pontos do espaço, suas localizações no espaço podem ser regulares ou irregulares e estas localizações podem compor, elas mesmas, um conjunto discreto ou contínuo.

⁴¹ Cressie cita o trabalho de Edmond Halley de 1686, "An historical account of trade winds, and monsoons, observable in the seas between and near the tropics; with an attempt to assign the physical cause of said winds", onde o conhecido astrônomo britânico traçou num mapa direções de ventos alísios e monções entre e próximos aos trópicos, tentando atribuir a elas uma causa física.

Poderíamos, entretanto, atribuir uma forma genérica a um modelo espacial como se segue: seja $\mathbf{s} \in \mathbb{R}^d$ uma localização genérica de dados num espaço euclidiano de d dimensões e suponha que o dado $\mathbf{Z}(\mathbf{s})$ na localização espacial \mathbf{s} seja uma variável aleatória. Faça \mathbf{s} variar sobre o conjunto $D \subset \mathbb{R}^d$ de modo a gerar o processo aleatório:

$$\{\mathbf{Z}(\mathbf{s}): \mathbf{s} \in D\} \quad (3.1)$$

Onde D é usualmente assumido como fixo, embora seja possível generalizar para o caso em que D é aleatório.

Este processo aleatório apresentado na equação (3.1) pode ser dividido em três grandes grupos, de acordo basicamente com o tipo de dado que se tem.

A Geoestatística pode ser considerada uma disciplina híbrida da engenharia de mineração, geologia, matemática e estatística, se caracteriza por reconhecer a variabilidade espacial tanto em larga escala como em pequena escala, modelando tanto tendência espacial como autocorrelação. A principal distinção entre os problemas do tipo Geoestatísticos e os demais encontrados em estatística espacial é a capacidade do índice \mathbf{s} de variar continuamente sobre um sobre um subconjunto de \mathbb{R}^d .

Há também os dados de padrões pontuais. Estes surgem quando a principal variável a ser analisada é a localização do evento. Frequentemente a primeira questão a ser respondida é se o padrão apresenta caráter espacial aleatório, agrupamento ou regularidade. Há a possibilidade de que os dados não sejam variáveis quantitativas, mas, por exemplo imagens, que, neste caso, requerem uma análise sofisticada de modo a serem utilizadas num modelo do tipo da equação (3.1), que se transformaria num modelo do tipo Boleano.

Finalmente, existem os dados em treliças. Este tipo de dado é o que apresenta maior importância para este trabalho porque é nele que se enquadram os métodos da econometria espacial. Uma "treliça" de locações vem da idéia de pontos do \mathbb{R}^d espaçados (regiões), ligados aos seus "vizinhos", Como o exemplo da figura abaixo:

1	2	3
4	5	6

figura 3.1

A região 1 é vizinha das regiões 2, 4 e 5, enquanto a região 5 é vizinha de todas as demais. Estas relações de vizinhança podem ser representadas pela matriz de conectividade **w** abaixo:

$$w = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Ou pela matriz **W**, que seria a matriz **w** normalizada pelas linhas, isto é, alterada de tal modo que a soma em cada linha seja exatamente igual a 1, isto é feito simplesmente dividindo o valor de cada elemento da matriz pelo total das linhas⁴².

$$W = \begin{bmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \end{bmatrix}$$

Considerando, em ambos os casos, vizinhança de ordem 1. A região 1 não é vizinha (de ordem 1) da região 3, mas elas são vizinhas de segunda ordem, pois a região 1 é vizinha da região 2 que por sua vez é vizinha da região 3. Para relações de ordem 2 ou superiores são necessárias, portanto, diferentes matrizes de conectividade. A matriz **W** de ordem 0 (**W₀**) é a própria matriz identidade.

⁴² V. Anselin(1988)

3.4.2 Autocorrelação espacial

O problema da autocorrelação espacial tem alguma semelhança com a autocorrelação temporal. De fato, se as regiões de um determinado espaço fossem todas "enfileiradas", de tal modo que só existisse o vizinho de "frente" e o de "trás", (ou, em termos estatísticos, só pudessem apresentar dependência unidirecional) como mostra a figura abaixo, recairíamos numa situação formalmente idêntica à das séries de tempo e, portanto, todo o tratamento econométrico seria idêntico ao das séries de tempo



figura 3.2

Um espaço como o da figura acima é, evidentemente, algo raro de se obter. O caso mais geral é como o da figura 3.2 (embora não necessariamente com a mesma regularidade), onde os dados (regiões) estão dispostos numa superfície bidimensional, apresentando dependência bidirecional.

O problema de autocorrelação espacial é, em princípio o mesmo da autocorrelação temporal. Num modelo de regressão, se os erros são correlacionados entre si (temporal ou espacialmente), os estimadores de mínimos quadrados ordinários serão ineficientes e os estimadores das variâncias serão viesados invalidando os testes de significância.

Todavia, há uma segunda razão pela qual a identificação da autocorrelação é importante, como nota Griffith(1987), que é conceitual em natureza, particularmente no contexto de modelos de interação espacial. Assim sendo, deve-se atentar para a presença de estruturas de autocorrelação para a correta especificação do modelo.

Os processos de autocorrelação espacial também guardam analogia com os de séries de tempo

Enquanto a situação de autocorrelação serial, de ordem 1, pode ser representada da seguinte forma:

$$z_t = \mu_t + \rho z_{t-1} \quad (3.2)$$

Onde μ_t é um ruído branco, e ρ é o coeficiente de correlação, a autocorrelação espacial, também de ordem 1, é mostrada abaixo:

$$\mathbf{z} = \mu + \rho \mathbf{W}_1 \mathbf{z} \quad (3.3)$$

Este processo é conhecido como SAR (*spatial autoregressive*), onde \mathbf{W}_1 é a matriz de conectividade de 1ª ordem. Este é, portanto, um SAR (1).

Mais genericamente, podemos ter um SARMA (*spatial autoregressive moving average*). Abaixo, mostramos um SARMA(1,1)

$$\mathbf{z} = \mu + \rho \mathbf{W}_1 \mathbf{z} + \theta \mathbf{W}_1 \mu \quad (3.4)$$

Que pode facilmente incluir ordens superiores, bastando para isso incluir as respectivas matrizes de conectividade. Por exemplo, o processo abaixo seria um SAR(2):

$$\mathbf{z} = \mu + \rho_1 \mathbf{W}_1 \mathbf{z} + \rho_2 \mathbf{W}_2 \mathbf{z} \quad (3.5)$$

O mais conhecido teste para autocorrelação espacial⁴³ é o I de Moran, cuja forma utilizando a matriz de conectividade normalizada por colunas é dado por:

$$I = \frac{\mathbf{z}' \mathbf{W}_1 \mathbf{z}}{\mathbf{z}' \mathbf{z}} \quad (3.6)$$

Valores altos para I, em módulo, representam a existência de autocorrelação. O sinal desta última corresponde ao próprio sinal de I.

Mais precisamente, mostra-se que o índice de Moran tem distribuição assintótica normal e média e variância dadas por⁴⁴:

$$E(I) = -1/(n-1) \quad (3.7)$$

$$\text{Var}(I) = (\text{tr}(\mathbf{M}\mathbf{W}\mathbf{M}\mathbf{W}') + \text{tr}(\mathbf{M}\mathbf{W})^2 + 1)/(n^2 - 1) - [E(I)]^2 \quad (3.8)$$

⁴³ Para uma discussão mais detalhada sobre este e outros testes veja Anselin e Florax(1995b).

⁴⁴ V. Anselin(1988) e Griffith(1987).

Onde \mathbf{M} é a matriz de projeção dada por $\mathbf{I}_n - \mathbf{1}(\mathbf{1}'\mathbf{1})\mathbf{1}'$, sendo \mathbf{I}_n a matriz identidade de ordem n e $\mathbf{1}$ um vetor de dimensões $n \times 1$, tal que $\mathbf{1}_i = 1$, $i = 1, \dots, n$. E "tr" é o operador traço, que dá a soma dos elementos da diagonal principal de uma matriz. A variável I^* definida abaixo:

$$I^* = \frac{I - E(I)}{[\text{var}(I)]^{1/2}} \quad (3.9)$$

Terá portanto, uma distribuição normal padronizada, com a qual se pode proceder o teste, onde a hipótese nula é a ausência de autocorrelação espacial.

Um outro teste, sugerido originalmente por Burridge(1980), identificado aqui por LM-ERR, é baseado no princípio do multiplicador de Lagrange, é dado por:

$$\text{LM-ERR} = \frac{(\mathbf{z}'\mathbf{W}\mathbf{z}/s^2)^2}{T} \quad (3.10)$$

Onde $s^2 = \mathbf{z}'\mathbf{z}/n$ e $T = \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)$. LM-ERR é distribuído como uma χ^2 com 1 grau de liberdade. A hipótese nula é a mesma, isto é, a ausência de autocorrelação.

3.4.3 Causalidade Espaço-Temporal

A existência de autocorrelação não mostra uma relação de causa e efeito. Por exemplo a instalação de um parque pode ocasionar a valorização dos imóveis de seus bairros vizinhos, o que é um efeito que ocorre de uma vez por todas. A questão é saber se, uma alteração numa variável neste contexto pode se **propagar** pelos seus vizinhos.

Por exemplo, o aumento da criminalidade em um bairro. Este evento, *per se*, poderia influenciar as taxas de criminalidade em bairros vizinhos, como prevê o modelo do capítulo 1.

Então, o problema econométrico a ser estudado é: pode o valor de uma variável no instante t ter sido causado pelo valor desta mesma

variável, em uma região vizinha, no instante $t-1$ (ou, mais genericamente, $t-k$)?

É aí que faz-se necessária uma adaptação do teste de causalidade de Granger às condições espaciais. Trata-se então, de uma relação de causalidade não só no tempo, como no teste tradicional, mas também no espaço.

Considerando uma matriz \mathbf{y} de dimensões $n \times T$, cujos elementos são os números de homicídios, onde n é o número de regiões (distritos policiais, neste caso) e T o número de observações no tempo.

Definição 1: dada a matriz $\mathbf{A}(p \times q)$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & a_{2q} \\ \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \dots & a_{pq} \end{bmatrix}$$

Chamamos $\text{Vec}(\mathbf{A})$ o vetor de dimensões $pq \times 1$ dado por:

$$\text{Vec}(\mathbf{A}) = \begin{bmatrix} a_{11} \\ a_{12} \\ \dots \\ a_{1q} \\ a_{21} \\ \dots \\ a_{2q} \\ \dots \\ a_{p1} \\ \dots \\ a_{pq} \end{bmatrix}$$

Desta forma, $\text{Vec}(\mathbf{y})$ é um vetor de dimensões $nT \times 1$.

Façamos $\mathbf{v} = \mathbf{W}\mathbf{y}$, onde \mathbf{W} é a matriz de conectividade normalizada pelas linhas, portanto de dimensão $n \times n$. Sendo \mathbf{y} de dimensão $n \times T$, \mathbf{v} terá também dimensão $n \times T$.

Desta forma, cada elemento de $\text{Vec}(\mathbf{v})$ é a média ponderada⁴⁵ das vizinhanças dos elementos correspondentes de $\text{Vec}(\mathbf{y})$.

O teste será feito utilizando-se as seguintes equações:

$$Y_{it} = \alpha_0 + \alpha_1 Y_{i,t-1} + \alpha_2 Y_{i,t-2} + \dots + \alpha_m Y_{i,t-m} + \beta_1 V_{i,t-1} + \beta_2 V_{i,t-2} + \dots + \beta_m V_{i,t-m} + \varepsilon_t \quad (3.11)$$

Onde temos um total de $n(T-m)$ observações. Comparamos esta equação com a equação restrita:

$$Y_{it} = \alpha_0 + \alpha_1 Y_{i,t-1} + \alpha_2 Y_{i,t-2} + \dots + \alpha_m Y_{i,t-m} + \varepsilon_t \quad (3.12)$$

E aí o teste recai num teste de Granger tradicional onde procura-se testar a hipótese nula "V não causa Y", o que equivale a testar a hipótese $\beta_1 = \beta_2 = \dots = \beta_m = 0$.

O teste será dado pela comparação da soma dos quadrados dos resíduos da equação restrita (SQR) com a mesma soma para a equação não restrita (SQNR). Obteremos a seguinte estatística:

$$F = (n(T-m) - 2m - 1)(SQR - SQNR) / (m \cdot SQNR) \quad (3.13)$$

Que segue uma distribuição $F(m, n(T-m) - 2m - 1)$.

3.4.4 Os modelos STARMA

A classe de modelos STARMA é uma extensão dos modelos ARMA (de Box e Jenkins), e expressa z_{it} , onde $i = 1, 2, \dots, N$ representa a região em que ela se localiza, como uma combinação linear de observações e erros passados que podem estar defasados tanto no tempo como no espaço. A junção dos valores defasados espacialmente se dá pela matriz de conectividade \mathbf{W} , podendo, em princípio, haver defasagens espaciais de ordens superiores a 1. Se isto ocorrer, utilizam-se diferentes especificações \mathbf{W}_l , onde l é a ordem de defasagem espacial.

Fazendo \mathbf{z}_t o vetor ($N \times 1$) de observações no período t , podemos expressar a classe de modelos STARMA como se segue:

⁴⁵ No caso geral. Da maneira como foi definida a matriz \mathbf{W} , é a média aritmética simples dos valores das vizinhanças.

$$z_t = \alpha + \sum_{k=1}^p \sum_{l=0}^{\lambda_k} \phi_{kl} \mathbf{W}_l z_{t-k} + \varepsilon_t - \sum_{k=1}^q \sum_{l=0}^{m_k} \theta_{kl} \mathbf{W}_l \varepsilon_{t-k} \quad (3.14)$$

Onde p e q são as ordens dos processos autorregressivos e de médias móveis, respectivamente; λ_k é a ordem espacial do k-ésimo termo autorregressivo; m_k é a ordem espacial do k-ésimo termo de média móvel; ϕ_{kl} e θ_{kl} são os parâmetros autorregressivos e de médias móveis com defasagem temporal k e defasagem espacial l.

A título de exemplo, o processo STARMA(2₁₀, 1₁) seria dado por:

$$z_t = \alpha + \phi_{10} z_{t-1} + \phi_{11} \mathbf{W}_1 z_{t-1} + \phi_{20} z_{t-2} + \varepsilon_t - \theta_{10} \varepsilon_{t-1} - \theta_{11} \mathbf{W}_1 \varepsilon_{t-1} \quad (3.15)$$

A identificação de um modelo STARMA, a exemplo dos modelos ARMA, se dá através das funções de autocorrelação e autocorrelação parcial. A diferença é que, num modelo STARMA, estas funções saem das covariâncias entre a variável e suas defasagens, tanto no espaço como no tempo.

A função de covariância espaço-temporal é dada por:

$$\gamma_{sk} = E \left[\frac{z_t' \mathbf{W}_k z_{t+s}}{N} \right] \quad (3.16)$$

Segue que a função de autocorrelação espaço temporal será dada por:

$$\rho_{sk} = \frac{\gamma_{sk}}{[\gamma_{0k} \gamma_{00}]^{1/2}} \quad (3.17)$$

Esta função pode ser estimada diretamente por:

$$\hat{\rho}_{0k}(s) = \frac{\sum_{i=1}^N \sum_{t=s}^T z_i L^k z_{i,t-s}}{\sum_{i=1}^N \sum_{t=1}^T z_i^2 \sum_{i=1}^N \sum_{t=1}^T (L^k z_i)^2} \quad (3.18)$$

Onde L^k é o operador defasagem espacial de ordem k, isto é, $L^k z_{it}$ é o elemento da linha i e coluna t da matriz $\mathbf{W}_k z$.

A função de autocorrelação parcial pode ser obtida diretamente do modelo genérico STAR(k, λ, \dots):

$$z_t = \sum_{j=1}^k \sum_{l=0}^{\lambda_k} \phi_{jl} W_l z_{t-j} + \varepsilon_t \quad (3.19)$$

A função de autocorrelação parcial pode ser obtida estimando-se sucessivamente a equação (3.19) para $l = 0, 1, \dots, \lambda$ e, para cada valor de l , $k = 0, 1, 2, \dots$. O último coeficiente obtido $\hat{\phi}_{kl}$ é a função de autocorrelação parcial para a defasagem espacial l e a defasagem temporal k .

Alternativamente, é possível obter as autocorrelações parciais através de um conjunto de equações análogas às equações tradicionais de Yule-Walker⁴⁶. Já que as equações de Yule-Walker (bem como suas análogas espaciais) necessitam dos valores das autocovariâncias, e estas por sua vez, têm seu valor real desconhecido, ficamos com as estimações da equação (3.19) por apresentarem um valor mais preciso.

A identificação dos modelos é semelhante aos de séries de tempo: os modelos STAR apresentam a função de autocorrelação caindo continuamente (tanto no espaço como no tempo) e a função de autocorrelação parcial cai abruptamente (trunca) na defasagem que corresponde à ordem do modelo; o contrário acontece com os modelos STMA, enquanto os modelos STARMA apresentam as duas funções caindo continuamente, como é resumido no quadro que se segue:

modelo	função de autocorrelação espaço-temporal	função de autocorrelação parcial espaço-temporal
STAR($p, \lambda_1, \dots, \lambda_p$)	cai continuamente	trunca em p defasagens temporais e λ_p defasagens espaciais
STMA(q, m_1, \dots, m_q)	trunca em q defasagens temporais e m_q defasagens espaciais	cai continuamente
STARMA(p, q)	cai continuamente	cai continuamente

⁴⁶ V. Pfeifer e Deutsch(1980a).

3.4.5 Outros desenvolvimentos em Econometria Espacial

Florax e Rey(1995) apresentam um estudo a respeito dos efeitos da má especificação de efeitos espaciais em modelos de regressão linear.

Com relação a modelos com variáveis qualitativas, McMillen(1995) estuda o efeito em modelos Probit, enquanto Dubin(1995) o faz para modelos Logit.

No que se refere a modelos que envolvem espaço e tempo, LeSage(1995) estima um modelo espaço-temporal de precificação de ativos.

4. Resultados Empíricos

4.1. Dados utilizados

Neste trabalho utilizaremos o número de homicídios ocorridos a cada mês na cidade de São Paulo, dividida pelas áreas correspondentes a 93 distritos policiais, por um período de dois anos (1995 e 1996). Teremos por tanto, um total de $93 \times 24 = 2232$ dados. Estes dados foram obtidos no Anuário do Departamento de Homicídios e Proteção à Pessoa da Polícia Civil do Estado de São Paulo.

Não se deve confundir esta divisão da cidade em distritos policiais com a divisão em distritos⁴⁷. A cidade é dividida em 8 seccionais de polícia: 1ª seccional (centro), 2ª e 6ª (zona sul), 5ª, 7ª e 8ª (zona leste), 3ª (zona oeste) e 4ª (zona norte). Cada seccional abrange um número de distritos policiais, que vai de 9 (8ª seccional) a 15 (3ª seccional).

Utilizaremos o número absoluto de homicídios. O número de homicídios *per capita*, não é viável neste caso, exatamente pelo fato de a divisão por distritos policiais não coincidir com a divisão por distritos, sendo esta última a que apresenta dados populacionais.

Este problema também ocorre quando queremos comparar, para cada distrito policial, o número de homicídios com a renda, pois esta informação também está disponível por distritos. Na tabela 4.1 procurou-se contornar esta dificuldade encontrando, para cada distrito policial, o distrito "correspondente", isto é, aquele que ocupa a maior parte da área do distrito policial em questão. A renda mostrada foi obtida pelo censo de 1991 (IBGE) e é a renda média, por distrito, dos chefes de família, medida em dólares.

⁴⁷ Utilizada, por exemplo, pelo IBGE no recenseamento.

Tabela 4.1: Homicídios por distrito policial e renda

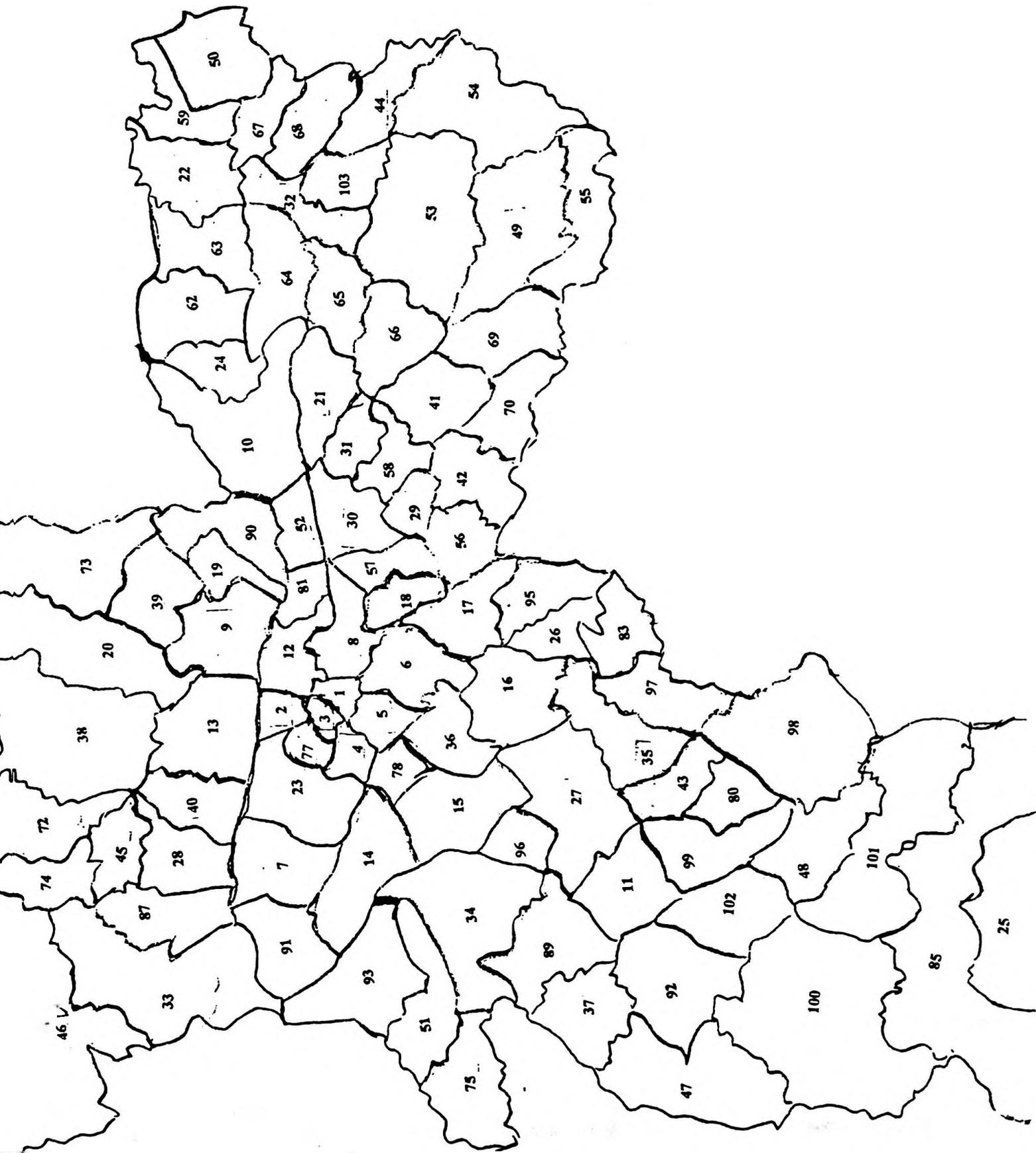
Distrito Policial	homicídios 1995	homicídios 1996	Distrito Correspondente	Renda (US\$)
1° - Sé	29	32	Sé	528,21
2° - Bom Retiro	17	17	Bom Retiro	571,19
3° - Santa Ifigênia	50	37	República	726,03
4° - Consolação	14	15	Consolação	1528,22
5° - Liberdade	43	38	Liberdade	962,94
6° - Cambuci	66	29	Cambuci	709,68
8° - Brás	39	41	Brás	556,32
12° - Pari	35	35	Pari	534,92
77° - Santa Cecília	37	50	Santa Cecília	946,43
78° - Paulista	2	5	Bela Vista	1127,78
16° - V.Clementino	29	34	Saúde	1107,40
17° - Ipiranga	31	31	Ipiranga	696,90
26° - Sacomã	32	71	Sacomã	544,63
27° - Ibirapuera	27	20	Moema	2033,36
35° - Jabaquara	78	161	Jabaquara	629,53
36° - Vila Mariana	14	11	Vila Mariana	1390,53
83° - Parque Bristol	39	10	Cursino	736,87
95° - Heliópolis	79	112	Sacomã	544,63
96° - Monções	23	13	Itaim Bibi	1565,26
97° - Americanópolis	57	31	Cidade Ademar	496,12
7° - Lapa	23	22	Lapa	897,59
14° - Pinheiros	20	25	Pinheiros	1678,60
15° - Jardim Paulista	26	20	Jardim Paulista	2074,78
23° - Perdizes	16	22	Perdizes	1430,84
33° - Pirituba	29	34	Pirituba	500,35
34° - Morumbi	38	61	Morumbi	2653,11
37° - Campo Limpo	134	179	Campo Limpo	469,47
46° - Perus	33	27	Perus	325,56
51° - Butantã	45	119	Butantã	1118,05
74° - Parada de Taipas	84	137	Jaraguá	403,01
75° - Jardim Arpoador	55	25	R. Tavares	451,66
87° - V. P. Barreto	46	45	Pirituba	500,35
89° - Jardim Taboão	53	19	Vila Andrade	1652,04
91° - CEASA	25	19	Vila Leopoldina	884,29
93° - Jaguaré	46	52	Jaguaré	721,91
9° - Vila Guilherme	21	24	Vila Guilherme	560,79
13° - Casa Verde	21	27	Santana	981,36
19° - Vila Maria	29	21	Vila Maria	1390,53
20° - Tucuruvi	25	29	Tucuruvi	655,11
28° - N.Senhora do Ó	28	40	Freguesia do Ó	505,20
38° - V.N.Cachoeirinha	63	112	Cachoeirinha	388,09
39° - Vila Gustavo	33	23	Vila Medeiros	416,09
40° - Bairro do Limão	37	45	Limão	491,34
45° - Vila Brasilândia	49	43	Brasilândia	326,47
72° - Vila Penteado	67	38	Brasilândia	326,47
73° - Jaçanã	66	79	Jaçanã	457,98
90° - Pq. Novo mundo	50	52	Vila Maria	1390,53

10° - Penha de França	36	35	Cangaíba	431,41
18° - Alto da Moóca	6	7	Moóca	876,53
21° - Vila Matilde	39	21	Vila Matilde	496,82
29° - Vila Diva	10	18	Água Rasa	583,14
30° - Tatuapé	10	11	Tatuapé	821,50
31° - Vila Carrão	10	6	Carrão	547,40
42° - Parque S. Lucas	24	9	Vila Prudente	546,63
52° - Parque S. Jorge	8	2	Tatuapé	821,50
56° - Vila Alpina	36	31	Vila Prudente	546,63
57° - Parque da Moóca	3	2	Moóca	876,53
58° - Vila Formosa	20	11	Vila Formosa	537,09
81° - Belém	14	25	Belém	593,09
11° - Santo Amaro	56	60	Santo Amaro	1652,56
25° - Parelheiros	82	86	Parelheiros	316,65
43° - Cidade Ademar	53	57	Cidade Ademar	496,12
47° - Capão Redondo	256	233	Capão Redondo	376,31
48° - Cidade Dutra	41	41	Cidade Dutra	501,90
80° - Vila Joaniza	49	64	Campo Grande	1013,87
85° - Jardim Mirna	103	74	Cidade Dutra	501,90
92° - Pq. Sto. Antônio	201	186	Jardim São Luís	421,39
98° - Jardim Miriam	148	140	Pedreira	398,25
99° - Campo Grande	35	22	Campo Grande	1013,87
100° - Jd. Herculano	150	156	Jardim Ângela	314,33
101° - Jd. das Imbulas	154	155	Grajaú	344,90
102° - Socorro	22	20	Socorro	837,37
22° - S.Miguel Paulista	102	119	Jardim Helena	262,00
24° - Ermel. Matarazzo	29	21	Cangaíba	431,41
32° - Itaquera	56	93	Itaquera	370,62
50° - Itaim Paulista	121	133	Itaim Paulista	275,28
59° - Jardim dos Ipês	85	25	Jardim Helena	262,00
62° - Jardim Popular	61	47	Erm. Matarazzo	376,36
63° - Vila Jacuí	64	27	Vila Jacuí	342,73
64° - Cid.A.E.Carvalho	41	53	Itaquera	370,62
65° - Artur Alvim	42	23	Artur Alvim	407,23
67° - Jardim Robru	70	31	Lajeado	265,23
68° - Lajeado	79	116	Lajeado	265,23
103° - Cohab II	35	34	José Bonifácio	369,11
41° - Vila Rica	71	63	Aricanduva	453,12
44° - Guaianazes	48	54	Guaianazes	306,44
49° - São Mateus	115	121	São Mateus	375,82
53° - Parque do Carmo	44	49	Pq. do Carmo	364,60
54° - Cid. Tiradentes	64	78	Cid. Tiradentes	287,56
55° - Parque S.Rafael	47	17	São Rafael	360,37
66° - Jd. Aricanduva	28	25	Aricanduva	453,12
69° - Teotônio Vilela	80	73	São Mateus	375,82
70° - Sapopemba	51	91	Aricanduva	453,12
Total da cidade	4802	4897		

Pelos valores mostrados na tabela 1, encontramos uma correlação entre o número de homicídios e a renda de -0,3622 para 1995 e -0,3051 para 1996. Estes resultados devem ser visto com ressalva, tendo em vista a não coincidência entre distritos e distritos policiais⁴⁸, mas é um indício de que a incidência de homicídios tende a ser maior em regiões de renda mais baixa.

⁴⁸ Enfatize-se que, mesmo quando há coincidência nos nomes, a área de abrangência não é a mesma.

Distritos Policiais do Município de São Paulo
instalados até 01/01/92



4.2. Resultados em séries de tempo

Agrupando os dados dos diversos distritos policiais obtemos uma série de tempo com 24 dados mensais⁴⁹ para toda a cidade de São Paulo que são mostrados na tabela 4.2:

Tabela 4.2: homicídios na cidade de São Paulo

mês/ano	homicídios
Janeiro/1995	421
Fevereiro/1995	397
Março/1995	442
Abril/1995	453
Maio/1995	431
Junho/1995	382
Julho/1995	396
Agosto/1995	374
Setembro/1995	364
Outubro/1995	337
Novembro/1995	375
Dezembro/1995	430
Janeiro/1996	480
Fevereiro/1996	466
Março/1996	471
Abril/1996	398
Maio/1996	383
Junho/1996	381
Julho/1996	422
Agosto/1996	396
Setembro/1996	377
Outubro/1996	379
Novembro/1996	349
Dezembro/1996	395

O correlograma desta variável é mostrado na tabela 4.3:

⁴⁹ Ressalve-se que, na dimensão temporal, a amostra de que dispomos é considerada pequena para o tratamento em séries de tempo. Portanto, os resultados obtidos devem ser considerados com a devida ressalva.

Tabela 4.3: correlograma para homicídios na cidade de São Paulo
valor crítico: 0,4082

autocorrelação	autocorrel. parcial	estatística Q	probabil.
0,617	0,617	10,338	0,001
0,192	-0,305	11,387	0,003
-0,227	-0,344	12,912	0,005
-0,358	0,037	16,916	0,002
-0,407	-0,220	22,347	0,000
-0,302	-0,072	25,507	0,000
-0,246	-0,206	27,722	0,000
-0,156	-0,148	28,674	0,000
-0,060	-0,044	28,823	0,001
0,096	-0,012	29,230	0,001
0,183	-0,072	30,844	0,001
0,214	-0,051	33,225	0,001

O correlograma da tabela 4.3 apresenta a função de autocorrelação declinando, enquanto a autocorrelação parcial é truncada na primeira defasagem, o que é um indício de um processo AR(1).

Os resultados da estimação são mostrados na tabela 4.4:

Tabela 4.4: estimação de modelo AR(1) para a variável homicídios

variável	coeficiente	erro padrão	estatística t	probabil.
Constante	401,5542	17,22099	23,31772	0,0000
AR(1)	0,619067	0,170509	3,630690	0,0016

Portanto, a equação estimada para a variável homicídios (denominada z_t) é:

$$z_t = 401,55 + 0,62 z_{t-1} + \varepsilon_t \quad (4.1)$$

(17,22) (0,17)

O coeficiente do termo autorregressivo é significativo a 1% e apresenta um valor positivo relativamente alto, evidenciando a existência de autocorrelação temporal importante na variável homicídios.

Este modelo gera um vetor de resíduos com as características de um ruído branco, com mostra a tabela 4.5:

Tabela 4.5: correlograma dos resíduos para a estimação do AR(1)

autocorrelação	autocorrel. Parcial	estatística Q	probabil.
0,196	0,196	1,0050	
0,014	-0,025	1,0106	0,315
-0,367	-0,380	4,8806	0,087
-0,163	-0,023	5,6807	0,128
-0,214	-0,189	7,1466	0,128
-0,037	-0,124	7,1920	0,207
-0,068	-0,131	7,3571	0,289
-0,019	-0,184	7,3709	0,391
-0,020	-0,124	7,3870	0,496
0,079	-0,068	7,6622	0,569
0,064	-0,108	7,8589	0,643
0,107	-0,032	8,4566	0,672

Pelo teste de Ljung-Box aceita-se a hipótese de ruído branco para os resíduos, assim sendo, considera-se válido o modelo AR(1) estimado.

O mesmo procedimento será feito a seguir com cada uma das oito seccionais. Antes disso, para melhor dimensionarmos a magnitude do problema em cada uma delas, a tabela 4.6 mostra a média mensal de homicídios a elas referentes:

tabela 4.6: médias mensais de homicídios por seccional

1ª seccional	26,3
2ª seccional	37,6
3ª seccional	61,6
4ª seccional	42,6
5ª seccional	16,4
6ª seccional	110,2
7ª seccional	62,8
8ª seccional	46,6

Tabela 4.7: correlograma para homicídios na 1ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
0,074	0,074	0,1500	0,699
-0,114	-0,120	0,5184	0,772
0,079	0,099	0,7033	0,872
-0,273	-0,312	3,0354	0,552
-0,155	-0,079	3,8245	0,575
0,070	0,005	3,9935	0,678
-0,003	0,013	3,9938	0,780
0,005	-0,049	3,9948	0,858
0,047	-0,029	4,0853	0,906
0,064	0,076	4,2694	0,934
-0,085	-0,096	4,6131	0,948
0,005	0,035	4,6144	0,970

A variável homicídios tem comportamento semelhante ao de um ruído branco na 1ª seccional (centro), portanto ela não se enquadra num modelo do tipo ARMA.

Tabela 4.8: correlograma para homicídios na 2ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
0,593	0,593	9,5501	0,002
0,317	-0,054	12,399	0,002
-0,061	-0,351	12,508	0,006
-0,188	0,007	13,614	0,009
-0,191	0,073	14,817	0,011
-0,118	-0,034	15,298	0,018
-0,196	-0,303	16,704	0,019
-0,209	-0,053	18,413	0,018
-0,156	0,173	19,420	0,022
-0,155	-0,215	20,498	0,025
-0,100	-0,168	20,982	0,034
-0,138	-0,085	21,969	0,038

A exemplo do que ocorreu para o modelo geral da cidade, sugere-se um AR(1). Os coeficientes estimados foram:

tabela 4.9: estimação de AR(1) para a 2ª seccional

variável	coeficiente	erro padrão	estatística t	probabil.
Constante	37,12183	4,955647	7,490813	0,0000
AR(1)	0,669094	0,177323	3,773313	0,0011

O coeficiente estimado para a 2ª seccional (localizada na zona sul) é muito parecido com aquele estimado para toda a cidade.

Tabela 4.10: correlograma dos resíduos para a 2ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
0,116	0,116	0,3494	
0,199	0,188	1,4371	0,231
-0,275	-0,331	3,6095	0,165
-0,171	-0,157	4,4933	0,213
-0,246	-0,096	6,4312	0,169
0,063	0,105	6,5649	0,255
-0,098	-0,156	6,9066	0,330
-0,067	-0,253	7,0809	0,421
0,072	0,183	7,2933	0,505
-0,026	-0,044	7,3232	0,604
0,046	-0,124	7,4246	0,685
-0,027	-0,084	7,4636	0,760

Confirma-se a hipótese de ruído branco para os resíduos.

Para a 3ª seccional (zona oeste), temos

Tabela 4.11: correlograma para homicídios na 3ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
0,181	0,181	0,8849	0,347
-0,085	-0,122	1,0904	0,580
0,179	0,229	2,0463	0,563
-0,053	-0,167	2,1352	0,711
-0,170	-0,077	3,0839	0,687
0,016	0,007	3,0931	0,797
0,068	0,075	3,2634	0,860
0,009	0,032	3,2667	0,917
0,141	0,134	4,0989	0,905
0,226	0,147	6,3853	0,782
0,093	0,062	6,8033	0,815
-0,264	-0,341	10,431	0,578

A exemplo do que ocorre com a 1ª, os homicídios na 3ª seccional também se comportam como ruídos brancos.

Tabela 4.12: correlograma para homicídios na 4ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
-0,340	-0,340	3,1415	0,076
0,341	0,255	6,4379	0,040
0,035	0,251	6,4733	0,091
0,002	-0,013	6,4735	0,166
-0,064	-0,210	6,6075	0,252
-0,025	-0,137	6,6284	0,357
-0,204	-0,204	8,1495	0,320
0,028	-0,016	8,1798	0,416
-0,322	-0,235	12,485	0,187
-0,009	-0,193	12,489	0,254
-0,207	-0,185	14,551	0,204
0,138	0,190	15,544	0,213

Embora a um nível de significância menor, também aceitamos a hipótese de ruído branco para a 4ª seccional (zona norte).

Tabela 4.13: correlograma para homicídios na 5ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
-0,110	-0,110	0,3277	0,567
-0,036	-0,049	0,3652	0,833
0,153	0,145	1,0564	0,788
-0,138	-0,111	1,6530	0,799
0,331	0,332	5,2467	0,387
0,027	0,054	5,2716	0,509
-0,139	-0,074	5,9760	0,543
0,225	0,128	7,9545	0,438
-0,069	0,000	8,1512	0,519
-0,063	-0,146	8,3258	0,597
-0,132	-0,284	9,1649	0,607
-0,156	-0,127	10,427	0,579

Também a 5ª seccional (na zona leste) apresenta-se como um ruído branco.

Tabela 4.14: correlograma para homicídios na 6ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
0,573	0,573	8,9150	0,003
0,073	-0,381	9,0658	0,011
-0,289	-0,223	11,544	0,009
-0,327	0,036	14,881	0,005
-0,368	-0,355	19,320	0,002
-0,376	-0,233	24,212	0,000
-0,207	0,085	25,790	0,001
0,106	0,056	26,229	0,001
0,289	-0,099	29,704	0,000
0,375	0,242	35,959	0,000
0,282	-0,013	39,771	0,000
0,117	-0,118	40,480	0,000

Desta vez, o modelo a ser estimado é um AR(2).

Tabela 4.15: estimação de AR(2) para a 6ª seccional

variável	coeficiente	erro padrão	estatística t	probabil.
Constante	109,7743	5,006357	21,92697	0,0000
AR(1)	0,958027	0,193922	4,940273	0,0001
AR(2)	-0,510865	0,192298	-2,656629	0,0156

Interessante notar o sinal negativo do coeficiente de AR(2), o que pode indicar a existência de um padrão cíclico para esta seccional, que também se localiza na zona sul.

Tabela 4.16: correlograma dos resíduos para a 6ª seccional

autocorrelação	autocorrel. Parcial	estatística Q	probabil.
-0.141	-0.141	0.5002	
-0.037	-0.058	0.5364	
-0.063	-0.078	0.6465	0.421
-0.051	-0.077	0.7229	0.697
-0.105	-0.137	1.0661	0.785
-0.157	-0.221	1.8838	0.757
-0.169	-0.292	2.8843	0.718
0.203	0.054	4.4446	0.617
-0.156	-0.240	5.4282	0.608
0.151	-0.003	6.4342	0.599
0.027	-0.071	6.4680	0.692
0.206	0.132	8.7153	0.559

Para a 7ª seccional (na zona leste), temos:

Tabela 4.17: correlograma para homicídios na 7ª seccional

autocorrelação	autocorrel. Parcial	estatística Q	probabil.
0,457	0,457	5,6653	0,017
0,223	0,018	7,0717	0,029
0,116	0,010	7,4727	0,058
-0,152	-0,268	8,1934	0,085
-0,188	-0,039	9,3587	0,096
-0,202	-0,079	10,775	0,096
-0,038	0,190	10,827	0,146
-0,140	-0,255	11,595	0,170
-0,046	0,106	11,684	0,232
0,069	-0,012	11,897	0,292
0,026	0,073	11,929	0,369
0,198	0,142	13,963	0,303

Novamente, o modelo indicado é um AR(1).

tabela 4.18: estimação de AR(1) para a 7ª seccional

variável	coeficiente	erro padrão	estatística t	probabil.
Constante	60,78630	5,047127	12,04374	0,0000
AR(1)	0,507049	0,194571	2,605985	0,0165

O coeficiente obtido foi menor, na média, do que, por exemplo, o do modelo da cidade como um todo ou o da 2ª seccional. Estatisticamente falando, entretanto, não se pode rejeitar a hipótese de que eles sejam iguais.

Tabela 4.19: correlograma dos resíduos para a 7ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
-0,049	-0,049	0,0629	
0,066	0,063	0,1807	0,671
0,094	0,101	0,4347	0,805
-0,250	-0,249	2,3239	0,508
-0,050	-0,090	2,4049	0,662
-0,222	-0,216	4,0664	0,540
0,247	0,321	6,2504	0,396
-0,148	-0,193	7,0929	0,419
0,022	0,016	7,1131	0,524
0,164	-0,015	8,3085	0,503
-0,212	-0,062	10,470	0,400
0,074	-0,028	10,759	0,464

Finalmente, para a 8ª seccional (também na zona leste), obtemos:

Tabela 4.20: correlograma para homicídios na 8ª seccional

autocorrelação	autocorrel. parcial	estatística Q	probabil.
0,328	0,328	2,9132	0,088
0,059	-0,054	3,0120	0,222
-0,122	-0,140	3,4543	0,327
-0,360	-0,313	7,5065	0,111
-0,470	-0,331	14,769	0,011
-0,175	0,035	15,831	0,015
0,105	0,149	16,237	0,023
-0,014	-0,285	16,244	0,039
0,040	-0,229	16,309	0,061
0,003	-0,216	16,310	0,091
-0,058	-0,036	16,469	0,125
-0,092	-0,078	16,914	0,153

Este é um modelo de difícil identificação, mesmo porque não é aceita a hipótese de ruído branco por muito pouco. O modelo que melhor se ajustou, entretanto, foi um MA(1).

tabela 4.21: estimação de MA(1) para a 8ª seccional

variável	coeficiente	erro padrão	estatística t	probabil.
Constante	46,56899	1,717688	27,11144	0,0000
MA(1)	0,353868	0,201385	1,757171	0,0928

Tabela 4.22: correlograma dos resíduos para a 8ª seccional

autocorrelação	autocorrel. Parcial	estatística Q	probabil.
0,006	0,006	0,0010	
0,078	0,078	0,1733	0,677
-0,067	-0,068	0,3048	0,859
-0,233	-0,240	1,9917	0,574
-0,370	-0,385	6,4955	0,165
-0,127	-0,159	7,0506	0,217
0,179	0,233	8,2277	0,222
-0,076	-0,138	8,4542	0,294
0,061	-0,252	8,6098	0,376
-0,006	-0,282	8,6112	0,474
-0,018	-0,069	8,6269	0,568
-0,176	-0,125	10,229	0,510

É interessante notar que as duas regionais mais violentas (2ª e 6ª) apresentaram modelos autorregressivos. Entre as mais violentas, a única que se comportou como ruído branco foi a 3ª seccional. Isto talvez se deva ao fato de abranger o maior número de distritos policiais (15), bem como toda zona oeste, incluindo distritos policiais bastante violentos, como Campo Limpo, e alguns com níveis bastante razoáveis como Lapa e Pinheiros.

Entre as menos violentas, a 2ª seccional foi aquela que apresentou um modelo autorregressivo. Há que se notar, entretanto, que a 2ª seccional está na zona sul e é, portanto, vizinha da 6ª, o que pode indicar a presença de efeitos espaciais⁵⁰.

4.3. Testes de autocorrelação espacial.

Para verificar existência de autocorrelação espacial, utilizaremos os dados anuais para cada distrito policial, como na tabela 4.1.

⁵⁰ Um estudo mais aprofundado da existência de efeitos espaciais utilizando a agregação de seccionais complica-se à medida em que há apenas 8 delas, o que torna a amostra pequena em termos espaciais também. Nas seções seguintes, analisaremos os efeitos espaciais em termos dos distritos policiais, que são 93 no total.

A estatística mais comumente utilizada para se testar a existência de autocorrelação espacial é o índice I de Moran, que é dado, considerando uma matriz **W** normalizada pelas linhas, por:

$$I = \frac{\mathbf{z}'\mathbf{W}\mathbf{z}}{\mathbf{z}'\mathbf{z}} \quad (4.2)$$

Onde $\mathbf{z} = \mathbf{y} - \bar{y}$ é o vetor de homicídios em cada distrito num determinado ano centralizado, de tal modo que sua média seja zero.

Os valores dos índices de Moran para cada ano são:

$$I_{95} = 0,5326$$

$$I_{96} = 0,3128$$

O índice de Moran tem distribuição assintótica normal. A variável I^* definida abaixo:

$$I^* = \frac{I - E(I)}{[\text{var}(I)]^{1/2}} \quad (4.3)$$

Terá portanto, uma distribuição normal padronizada. Os valores encontrados foram:

$$I^*_{95} = 8,4958$$

$$I^*_{96} = 5,1449$$

Em ambos os casos, rejeitando-se a hipótese nula de ausência de correlação espacial, ao nível de significância de 1%, tendo em vista que o valor crítico da distribuição normal padronizada para 1% é de 2,57. Ademais, os valores positivos encontrados para I_{95} e I_{96} indicam que os homicídios são positivamente correlacionados no espaço.

O teste LM-ERR sugerido por Burridge(1980), é dado por:

$$\text{LM-ERR} = \frac{(\mathbf{z}'\mathbf{W}\mathbf{z}/s^2)^2}{T} \quad (4.4)$$

Onde $s^2 = z'z/n$ e $T = \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)$. LM-ERR é distribuído como uma χ^2 com 1 grau de liberdade.

Os valores encontrados foram:

$$\text{LM-ERR}_{95} = 65,4318$$

$$\text{LM-ERR}_{96} = 23,3621$$

O valor crítico da distribuição χ^2 com 1 grau de liberdade para o nível de significância de 1% (bicaudal) é 7,88. Portanto, também pelo teste LM-ERR rejeita-se a hipótese nula de ausência de autocorrelação espacial.

4.4 Teste de causalidade no espaço-tempo

Compararemos as duas equações:

$$y_{it} = \alpha_0 + \alpha_1 y_{i,t-1} + \alpha_2 y_{i,t-2} + \dots + \alpha_m y_{i,t-m} + \beta_1 v_{i,t-1} + \beta_2 v_{i,t-2} + \dots + \beta_m v_{i,t-m} + \varepsilon_t \quad (4.5)$$

$$y_{it} = \alpha_0 + \alpha_1 y_{i,t-1} + \alpha_2 y_{i,t-2} + \dots + \alpha_m y_{i,t-m} + \varepsilon_t \quad (4.6)$$

Com um total de $n(T-m)$ observações.

Testaremos a hipótese $\beta_1 = \beta_2 = \dots = \beta_m = 0$.

O teste será dado, como foi dito no capítulo 3, pelo comparação da soma dos quadrados dos resíduos da equação restrita (SQR) com a mesma soma para a equação não restrita (SQNR).

Pindyck e Rubinfeld(1998) sugerem que o teste seja feito para diferentes valores de m para verificar se o mesmo não é sensível para o número de defasagens. Os resultados estão na tabela 6:

Tabela 4.23: resultados dos testes de causalidade.

nº de defasagens (m)	estatística F	probabilidade
1	101,03	0,0000
2	20,51	0,0000
3	8,27	0,0000
4	3,57	0,0066

A hipótese de não causalidade (que poderíamos chamar aqui de não propagação) é rejeitada para todos os valores de m utilizados. Há indícios, portanto, de que o crime (de homicídio, no caso específico) se propaga ao longo do tempo entre os vizinhos.

4.5. O Modelo STARMA

Nesta seção procederemos a estimação do modelo STARMA (*space-time autorregressive moving average*). O procedimento adotado é o mesmo utilizado por Pfeifer e Deutsch(1980a e 1980b).

4.5.1 Identificação

Os valores das funções de autocorrelação e autocorrelação parcial espaço-temporais são dados na tabelas 4.24 e 4.25:

Tabela 4.24: função de autocorrelação espaço-temporal (valores críticos = $\pm 0,4082$)

defasagens temporais	defasagens espaciais		
	0	1	2
1	0,6972	0,4525	0,2870
2	0,6488	0,4135	0,2563
3	0,5933	0,3852	0,2267
4	0,5481	0,3648	0,2087
5	0,5198	0,3398	0,1933
6	0,4852	0,3255	0,1879
7	0,4674	0,3121	0,1905
8	0,4508	0,3114	0,1791
9	0,4252	0,3117	0,1764
10	0,4028	0,2934	0,1753
11	0,3578	0,2765	0,1617
12	0,3113	0,2520	0,1414

Tabela 4.25: função de autocorrelação parcial espaço-temporal (valores críticos = $\pm 0,4082$)

defasagens temporais	defasagens espaciais		
	0	1	2
1	0,7307	0,2585	0,0686
2	0,4579	0,1612	0,1208
3	0,3769	0,1423	0,1841
4	0,3403	0,0982	0,2413
5	0,3232	0,0992	0,3065
6	0,3156	0,0646	0,2692
7	0,3018	0,0581	0,2942
8	0,2962	0,0338	0,2382
9	0,2896	-0,0042	0,2509
10	0,2864	-0,0270	0,1975
11	0,2962	-0,0324	0,1664
12	0,2874	-0,0306	0,2187

A escolha do número de ordens espaciais é arbitrária. Como dizem Pfeifer e Deutsch(1980a): "é importante que seja tão grande quanto a ordem máxima de qualquer modelo que se tenha por hipótese". Nos limitamos nas tabelas acima à ordem 2, o que parece suficiente tendo em vista os resultados obtidos.

A função de autocorrelação parece cair continuamente tanto no espaço como no tempo, enquanto a função de autocorrelação parcial é truncada, o que é sugestivo de um modelo STAR.

4.5.2 Estimação e diagnóstico

Como a função de autocorrelação parcial é truncada na defasagem de ordem 2, mas não atinge o mesmo valor no espaço (a rigor, nem a de ordem 1 foi significativa), o modelo a ser tentado é um STAR(2₁₀), mostrado abaixo:

$$z_t = \alpha + \phi_{10} z_{t-1} + \phi_{11} W_1 z_{t-1} + \phi_{20} z_{t-2} + \varepsilon_t \quad (4.7)$$

Note-se que este é um modelo na forma linear tradicional $Y = X\beta + e$. O vetor de parâmetros dados por mínimos quadrados ordinários, dados por $\Phi = (X'X)^{-1}(X'Y)$ é, neste caso, a estimação por mínimos quadrados condicionais dos parâmetros do modelo STAR(2₁₀).

É preciso notar, no entanto, que as hipóteses usuais do método de mínimos quadrados não são obedecidas aqui: primeiramente, \mathbf{X} é uma matriz estocástica, e não fixa; ademais, não há independência entre os componentes da matriz \mathbf{X} e o erro. Mas é possível mostrar que as propriedades do modelo de regressão linear clássico são assintoticamente válidas para modelos com as características de um STAR⁵¹. Devemos lembrar que os resultados da estimação por mínimos quadrados ordinários são apenas aproximadamente válidos.

Os resultados das estimações são dados na tabela 4.26:

Tabela 4.26: resultados das estimações do modelo STAR(2₁₀)

variável	coeficiente	erro padrão	estatística t
constante	0,2648	0,1133	2,33
z_{t-1}	0,4281	0,0207	20,68
z_{t-2}	0,3483	0,0206	16,91
$\mathbf{W} z_{t-1}$	0,1601	0,0250	6,40

$$\sigma^2 = 7,6015$$

O modelo estimado foi, portanto:

$$\hat{z}_t = 0,26 + 0,43 z_{t-1} + 0,16 \mathbf{W}_1 z_{t-1} + 0,35 z_{t-2} \quad (4.8)$$

Para saber se este modelo é adequado, de maneira semelhante àquela usual nos modelos de Box e Jenkins, devemos verificar se os resíduos da equação se comportam como um ruído branco. Para isso, recorreremos de novo às funções de autocorrelação e autocorrelação parcial espaço-temporais, desta feita para os resíduos, cujos resultados estão nas tabelas a seguir:

⁵¹ V. Pfeifer e Deutsch(1980a).

Tabela 4.27: função de autocorrelação dos resíduos
(valores críticos = $\pm 0,4264$)

defasagens temporais	defasagens espaciais		
	0	1	2
1	-0,0677	-0,0158	0,0499
2	-0,1574	-0,0604	0,0050
3	0,0185	-0,0274	-0,0224
4	0,0219	-0,0095	-0,0255
5	0,0462	-0,0210	-0,0515
6	0,0248	-0,0403	-0,0106
7	0,0417	-0,0388	0,0340
8	0,0538	-0,0054	-0,0224
9	0,0265	0,0193	-0,0132
10	0,0728	0,0023	0,0322
11	0,0423	0,0305	0,0154
12	-0,0016	0,0248	-0,0113

Tabela 4.28: função de autocorrelação parcial dos resíduos
(valores críticos = $\pm 0,4264$)

defasagens temporais	defasagens espaciais		
	0	1	2
1	-0,0703	-0,0214	0,1512
2	-0,0857	-0,0563	0,2345
3	-0,0806	-0,0723	0,2727
4	-0,0712	-0,1031	0,2487
5	-0,0687	-0,1012	0,2610
6	-0,0597	-0,1030	0,2266
7	-0,0565	-0,1002	0,2329
8	-0,0563	-0,0769	0,1991
9	-0,0498	-0,0668	0,2335
10	-0,0696	-0,0646	0,2981
11	-0,1074	-0,1060	0,2129
12	-0,1172	-0,2130	0,2166

Pelos resultados das tabelas 11 e 12, podemos concluir que os resíduos se comportam, de fato, como um ruído branco e, assim sendo, o modelo estimado pode ser considerado adequado.

4.6. Simulações através do modelo STAR

Nesta seção simularemos, a partir do modelo STAR(2₁₀) estimado, a evolução da criminalidade nos dez anos seguintes ao último dado utilizado (dezembro de 1996). Para isto, utilizaremos parâmetros em torno dos valores estimados.

A tabela 4.29 resume a situação de dezembro de 1996:

tabela 4.29: situação real de dezembro de 1996

média de homicídios por distrito	4,2
desvio-padrão	4,6
número máximo por distrito	21
número mínimo por distrito	0
percentual entre 0 e 2 homicídios	48,4
percentual entre 3 e 5 homicídios	22,6
percentual entre 6 e 10 homicídios	19,4
percentual maior do que 10 homicídios	9,7

A tabela 4.30⁵² é o resultado da previsão utilizando os parâmetros estimados:

tabela 4.30: situação prevista de dezembro de 2006
 $\alpha = 0,26$ $\phi_{10} = 0,43$ $\phi_{11} = 0,16$ $\phi_{20} = 0,35$ $\sigma^2 = 7,6$

média de homicídios por distrito	6,2
desvio-padrão	3,7
número máximo por distrito	16
número mínimo por distrito	0
percentual entre 0 e 2 homicídios	18,06
percentual entre 3 e 5 homicídios	27,42
percentual entre 6 e 10 homicídios	40,32
percentual maior do que 10 homicídios	14,19

A média aumenta e o desvio-padrão se reduz, o que é compatível com a hipótese de espalhamento dos crimes pela cidade acompanhando o aumento da criminalidade. Reduz-se drasticamente o número de distritos policiais onde há 2 ou menos homicídios no mês.

Nas tabelas seguintes apresentaremos os resultados obtidos com modificações no parâmetro espaço-temporal ϕ_{11} .

⁵² Os resultados das tabelas de 4.30 em diante foram obtidos através de 10 simulações com os parâmetros indicados.

tabela 4.31: situação simulada de dezembro de 2006
 $\alpha = 0,26$ $\phi_{10} = 0,43$ $\phi_{11} = 0,2$ $\phi_{20} = 0,35$ $\sigma^2 = 7,6$

média de homicídios por distrito	12,6
desvio-padrão	4,3
número máximo por distrito	22
número mínimo por distrito	1
percentual entre 0 e 2 homicídios	1,94
percentual entre 3 e 5 homicídios	4,30
percentual entre 6 e 10 homicídios	22,58
percentual maior do que 10 homicídios	71,18

tabela 4.32: situação simulada de dezembro de 2006
 $\alpha = 0,26$ $\phi_{10} = 0,43$ $\phi_{11} = 0,1$ $\phi_{20} = 0,35$ $\sigma^2 = 7,6$

média de homicídios por distrito	4,1
desvio-padrão	3,2
número máximo por distrito	13
número mínimo por distrito	0
percentual entre 0 e 2 homicídios	34,62
percentual entre 3 e 5 homicídios	33,12
percentual entre 6 e 10 homicídios	29,25
percentual maior do que 10 homicídios	3,01

Podemos observar que tanto com o aumento de ϕ_{11} (tabela 4.31), como com sua queda (4.32), ocorrem mudanças importantes no resultado final, tanto na média como na distribuição dos homicídios pelos distritos.

Quando aumentamos ϕ_{11} para 0,2, a média pula para 12 homicídios por mês por distrito policial e há cerca de 70% dos distritos policiais superando a marca dos 10 homicídios mensais.

Ao reduzir para 0,1, a média cai para 4 e os distritos policiais com 2 ou menos homicídios mensais constituem maioria relativa (cerca de 35%).

Estes são fortes indicadores de que o componente espacial é relevante. Na tabela seguinte, deixamos ϕ_{11} igual a zero, isto é, retiramos por completo o componente espacial.

tabela 4.33: situação simulada de dezembro de 2006

$$\alpha = 0,26 \quad \phi_{10} = 0,43 \quad \phi_{11} = 0 \quad \phi_{20} = 0,35 \quad \sigma^2 = 7,6$$

média de homicídios por distrito	3,0
desvio-padrão	2,8
número máximo por distrito	11
número mínimo por distrito	0
percentual entre 0 e 2 homicídios	50,97
percentual entre 3 e 5 homicídios	28,39
percentual entre 6 e 10 homicídios	20,00
percentual maior do que 10 homicídios	0,65

O efeito final foi uma redução ainda maior na média e cerca de metade dos distritos policiais com 2 ou menos homicídios ao mês.

Nas tabelas seguintes o parâmetro a ser alterado é o autoregressivo "puro", ϕ_{10} .

tabela 4.34: situação simulada de dezembro de 2006

$$\alpha = 0,26 \quad \phi_{10} = 0,47 \quad \phi_{11} = 0,16 \quad \phi_{20} = 0,35 \quad \sigma^2 = 7,6$$

média de homicídios por distrito	12,4
desvio-padrão	4,7
número máximo por distrito	24
número mínimo por distrito	1
percentual entre 0 e 2 homicídios	2,37
percentual entre 3 e 5 homicídios	6,24
percentual entre 6 e 10 homicídios	25,38
percentual maior do que 10 homicídios	66,02

tabela 4.35: situação simulada de dezembro de 2006

$$\alpha = 0,26 \quad \phi_{10} = 0,39 \quad \phi_{11} = 0,16 \quad \phi_{20} = 0,35 \quad \sigma^2 = 7,6$$

média de homicídios por distrito	4,3
desvio-padrão	3,0
número máximo por distrito	12
número mínimo por distrito	0
percentual entre 0 e 2 homicídios	32,26
percentual entre 3 e 5 homicídios	36,13
percentual entre 6 e 10 homicídios	27,74
percentual maior do que 10 homicídios	3,87

Os efeitos também se mostram importantes. Fazendo uma comparação, entretanto, enquanto que os efeitos sobre a média são semelhantes, os efeitos sobre o desvio-padrão são maiores quando se altera o componente espaço-temporal, o que era esperado, tendo em vista o efeito de espalhamento que este último representa.

Nas tabelas 4.36 e 4.37, faremos um aumento de ϕ_{11} compensado por uma queda igual em ϕ_{10} , e vice-versa.

tabela 4.36: situação simulada de dezembro de 2006
 $\alpha = 0,26$ $\phi_{10} = \mathbf{0,39}$ $\phi_{11} = \mathbf{0,2}$ $\phi_{20} = 0,35$ $\sigma^2 = 7,6$

média de homicídios por distrito	6,2
desvio-padrão	3,5
número máximo por distrito	15
número mínimo por distrito	0
percentual entre 0 e 2 homicídios	19,14
percentual entre 3 e 5 homicídios	25,38
percentual entre 6 e 10 homicídios	43,44
percentual maior do que 10 homicídios	12,04

tabela 4.37: situação simulada de dezembro de 2006
 $\alpha = 0,26$ $\phi_{10} = \mathbf{0,47}$ $\phi_{11} = \mathbf{0,12}$ $\phi_{20} = 0,35$ $\sigma^2 = 7,6$

média de homicídios por distrito	5,7
desvio-padrão	3,9
número máximo por distrito	16
número mínimo por distrito	12
percentual entre 0 e 2 homicídios	24,30
percentual entre 3 e 5 homicídios	26,24
percentual entre 6 e 10 homicídios	37,63
percentual maior do que 10 homicídios	11,83

Quando aumento se deu no componente espaço-temporal, o efeito sobre a criminalidade foi ligeiramente maior, o que é um indício, ainda que leve, em se tratando da diferença obtida, que os efeitos espaciais têm maior influência na criminalidade. A diferença mais significativa, porém, se dá justamente na distribuição pelos distritos policiais.

As tabelas 4.38 e 4.39 se referem a um aumento do coeficiente espaço-temporal compensado por uma queda combinada dos dois coeficientes autorregressivos "puros" e vice-versa.

tabela 4.38: situação simulada de dezembro de 2006
 $\alpha = 0,26$ $\phi_{10} = \mathbf{0,41}$ $\phi_{11} = \mathbf{0,2}$ $\phi_{20} = \mathbf{0,33}$ $\sigma^2 = 7,6$

média de homicídios por distrito	5,8
desvio-padrão	3,6
número máximo por distrito	16
número mínimo por distrito	0
percentual entre 0 e 2 homicídios	21,72
percentual entre 3 e 5 homicídios	28,17
percentual entre 6 e 10 homicídios	38,28
percentual maior do que 10 homicídios	11,83

tabela 4.39: situação simulada de dezembro de 2006
 $\alpha = 0,26$ $\phi_{10} = 0,45$ $\phi_{11} = 0,12$ $\phi_{20} = 0,37$ $\sigma^2 = 7,6$

média de homicídios por distrito	6,0
desvio-padrão	3,8
número máximo por distrito	16
número mínimo por distrito	0
percentual entre 0 e 2 homicídios	21,08
percentual entre 3 e 5 homicídios	25,38
percentual entre 6 e 10 homicídios	40,43
percentual maior do que 10 homicídios	13,12

As diferenças apresentadas foram muito pequenas entre uma e outra situação.

Na tabela 4.40 é mostrada a simulação com os três parâmetros ligeiramente alterados de modo a, propositadamente, apresentar soma igual a 1 :

tabela 4.40: situação simulada de dezembro de 2006
 $\alpha = 0,26$ $\phi_{10} = 0,5$ $\phi_{11} = 0,2$ $\phi_{20} = 0,3$ $\sigma^2 = 7,6$

média de homicídios por distrito	28,8
desvio-padrão	5,5
número máximo por distrito	42
número mínimo por distrito	16
percentual entre 0 e 2 homicídios	0
percentual entre 3 e 5 homicídios	0
percentual entre 6 e 10 homicídios	0
percentual maior do que 10 homicídios	100

Com estes parâmetros, o comportamento é explosivo. Como a diferença em relação aos parâmetros inicialmente estimados não é grande, evidencia-se que esta situação não é uma realidade distante.

5. Conclusões

Após um histórico de modelos de “oferta” de crime, onde o número de crimes que o indivíduo (e, no agregado, o número de crimes oferecidos pelo “mercado”) é função de variáveis como a desutilidade da punição prevista pelo crime e a probabilidade de que esta punição efetivamente aconteça.

Exatamente a segunda, a probabilidade de aprisionamento, é utilizada num modelo em que, assumindo que ela seja função decrescente do próprio número de crimes, prevê a existência de causalidade, tanto no espaço como no tempo, da quantidade de crimes.

Por todos os modelos estimados, sejam eles de espaço-tempo, sejam simplesmente e tempo ou de espaço, verifica-se que a variável número de homicídios se propaga pelas dimensões temporal e espacial na cidade de São Paulo.

Nos modelos tradicionais, foi verificada a existência de um processo AR(1) bem configurado, ressalvado o problema do tamanho da amostra, no que se refere à série de tempo para a cidade como um todo. Quando desagregamos a série em seccionais, as mais violentas apresentaram modelos autorregressivos, enquanto as menos violentas se comportaram como um ruído branco, com exceção da 2ª seccional, que é vizinha da 6ª, a mais violenta, o que indica a presença de efeitos espaciais.

Os testes de autocorrelação espacial também foram inequívocos quanto à existência da mesma, mostrando a relação existente entre os crimes de homicídio nas diversas regiões abrangidas pelos distritos policiais.

A constatação da propagação indica que a política de segurança pública deve estar atenta a aumentos de criminalidade em regiões, procurando imediatamente combater este aumento reforçando a segurança tanto na própria região **como em regiões vizinhas**, no intuito de impedir a propagação no seu nascedouro. O combate intensificado apenas na região em questão pode ser ineficaz por ignorar os efeitos de propagação.

Fazendo simulações com valores em torno de parâmetros estimados de um modelo STARMA (mais precisamente um STAR), verificamos a tendência ao aumento da criminalidade, bem como seu espalhamento pela cidade. Para esta tendência, tanto o fator espaço-temporal como o fator temporal "puro" se mostraram importantes.

Além disso, percebe-se a necessidade de incremento importante no combate à criminalidade na cidade como um todo, tendo em vista à proximidade de uma situação explosiva. Importante lembrar que esta proximidade foi observada para dados de 1995 e 1996.

No que se refere aos modelos econométricos espaço-temporais estimados, o modelo STARMA tem a vantagem de apresentar parâmetros, permitindo assim uma maior análise e a simulação. Por sua vez, o teste de Causalidade de Granger Espaço-Temporal serve para constatar a propagação (o que pode servir para nortear políticas que visem a combatê-la ou estimulá-la), mas é obtido com um esforço computacional consideravelmente menor do que o modelo STARMA.

Bibliografia

- Abraham, B. (1983) "The Exact Likelihood Function for a Space-Time Model", *Metrika*, 30, 239-43.
- Alper, Neil O. e Daryl A. Hellman (editores) (1997) **Economics of Crime: A Reader**, 2ª ed., Simon & Schuster, Needham Heights.
- Anselin, Luc (1982) "Implicit Functional Relationships Between Systemic Effects in a General Model of Movement", *Regional Science and Urban Economics*, vol.12, nº 3, agosto.
- Anselin, Luc (1988) **Spatial Econometrics: Methods and Models**, Kluwer Academic Publishers, Dordrecht.
- Anselin, Luc (1990) "Some Robust Approaches to Testing and Estimation in Spatial Econometrics", *Regional Science and Urban Economics*, vol.20, nº 2, setembro.
- Anselin, Luc (1992) "Space and Applied Econometrics: Introduction", *Regional Science and Urban Economics*, vol.22, nº 3, setembro.
- Anselin, Luc e Raymond J.G.M. Florax (editores) (1995) **New Directions in Spatial Econometrics**, Springer, Berlim.
- Anselin, Luc e Raymond J.G.M. Florax (1995b) "Small Sample Properties of Tests for Spatial Dependence in Regression Models: Some Further Results", in: L. Anselin e R.J.G.M. Florax (editores) **New Directions in Spatial Econometrics**, Springer, Berlim, 21-74.
- Anselin, Luc, A. Varga e Z. Acs (1997) "Local Geographic Spillovers Between University Research and High Technology Innovations", *Journal of Urban Economics*, 42, nº 3.
- Balkin, Steven e John F. McDonald (1981) "The Market for Street Crime: An Economic Analysis of Victim-Offender Interaction", *Journal of Urban Economics*, 10, 390-405.
- Becker, Gary S. (1968) "Crime and Punishment: An Economic Approach", *Journal of Political Economy*, vol.76, nº 2, 169-217.
- Block, Michael K. e John M. Heineke (1975) "A Labor Theoretic Analysis of the Criminal Choice", *American Economic Review*, vol.65, nº 3, junho, 315-25.
- Burridge, P. (1980) "On the Cliff-Ord Test for Spatial Autocorrelation", *Journal of the Royal Statistical Society*, 42, 107-8.
- Cameron, Samuel (1988) "The Economics of Crime Deterrence: A Survey of Theory and Evidence", *Kyklos*, vol.41, fasc.2, 301-23.
- Chamberlain, E.H. (1933) **The Theory of Monopolistic Competition**, Cambridge.

- Cobb, William E. (1973) "Theft and the Two Hypotheses", in: **The Economics of Crime and Punishment**, American Enterprise Institute, Washington, 19-30.
- Cook, Philip J. e Gary A. Zarkin (1997) "Crime and Business Cycle" in: N.O. Alper. e D.A. Hellman (editores) **Economics of Crime: A Reader**, 2ª ed., Simon & Schuster, Needham Heights, 60-74.
- Cressie, Noel A.C. (1993) **Statistics for Spatial Data**, John Wiley & Sons, Nova York.
- Davidson, Rusell e James G. McKinnon (1993) **Estimation and Inference in Econometrics**, Oxford university Press, Nova York.
- Dilulio Jr., John J. (1996) "Help Wanted: Economists, Crime and Public Policy", *Journal of Economic Perspectives*, vol.10, nº 1, 3-24.
- Doratioto, Francisco (1991) **A Guerra do Paraguai**, Brasiliense, São Paulo.
- Dubin, Robin (1992) "Spatial Autocorrelation and Neighborhood Quality", *Regional Science and Urban Economics*, vol.22, nº 3, setembro.
- Dubin, Robin (1995) "Estimating Logit Models with Spatial Dependence", in: L.Anselin e R.J.G.M. Florax (editores) **New Directions in Spatial Econometrics**, Springer, Berlim, 229-42.
- Ehrlich, Isaac (1973) "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation", *Journal of Political Economy*, vol.81, nº 3, 521-65.
- Ehrlich, Isaac (1975) "The Deterrent Effect of Capital Punishment: A Question of Life and Death", *American Economic Review*, vol.65, nº 3, junho, 397-417.
- Ehrlich, Isaac (1977) "Capital Punishment and Deterrence: Some Further Thoughts and Additional Evidence", *Journal of Political Economy*, vol.85, nº 4, 741-88.
- Ehrlich, Isaac (1981) "On the Usefulness of Controlling Individuals: An Economic Analysis of Rehabilitation, Incapacitation and Deterrence", *American Economic Review*, vol.71, nº 3, junho, 307-22.
- Ehrlich, Isaac (1996) "Crime, Punishment and the Market for Offenses", *Journal of Economic Perspectives*, vol.10, nº 1, 43-67
- Ehrlich, Isaac e Gary Becker (1972) "Market Insurance, Self-Insurance and Self-Protection", *Journal of Political Economy*, vol.80, nº 4, 623-48..
- Ehrlich, Isaac e George Brower (1987) "On the Issue of Causality in the Economic Model of Crime and Law Enforcement: Some Theoretical Considerations and Experimental Evidence", *American Economic Review*, vol.77, nº 2, maio,99-110.
- Enders, Walter (1995) **Applied Econometric Time Series**, John Wiley & Sons, Nova York.

- Fender, John (1999), "A General Equilibrium Model of Crime and Punishment", *Journal of Economic Behavior & Organization*, vol. 39, 437-53.
- Fischer, Manfred M. e Arthur Getis (editores) (1997) **Recent Developments in Spatial Analysis**, Springer, Berlim.
- Florax, Raymond e H. Folmer (1992) "Specification and Estimation of Spatial Linear Regression Models: Monte Carlo Evaluation of Pre-Test Estimators", *Regional Science and Urban Economics*, vol.22, nº 3, setembro.
- Florax, Raymond J.G.M. e Serge Rey (1995) "The Impacts of Misspecified Spatial Interaction in Linear Regression Models", in: L. Anselin e R.J.G.M. Florax (editores) **New Directions in Spatial Econometrics**, Springer, Berlim, 111-35.
- Freitas, Eduardo F. (1993) **Análise de Causalidade Entre Preços e Oferta Monetária no Brasil**, dissertação de mestrado, mimeo, FEA/USP, São Paulo.
- Freeman, Richard B. (1996) "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?" *Journal of Economic Perspectives*, vol.10, nº 1, 25-42.
- Freeman, Scott, Jeffrey Grooger e Jon Sonstelie (1996) "The Spatial concentration of Crime", *Journal of Urban Economics*, 40, 216-31.
- Fujita, Masahisa, Paul Krugman e Anthony J. Venables (1999) **The Spatial Economy: Cities, Regions and International Trade**, MIT Press, Cambridge.
- Gaviria, Alejandro (2000) "Increasing Returns and the Evolution of Violent Crime: the Case of Colombia", *Journal of Development Economics*, vol.61, 1-25.
- Glaeser, Edward, Bruno Sacerdote e José Scheinkman (1996) "Crime and Social Interaction", *Quarterly Journal of Economics*, 111, 507-48.
- Georgescu-Roegen, Nicholas (1971) **The Entropy Law and the Economic Process**, Harvard University Press, Cambridge.
- Granger, C.W.J. (1969) "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", *Econometrica*, 36, 424-38.
- Granger, C.W.J. (1980) "Testing for Causality: A Personal Viewpoint", *Journal of Economic Dynamics and Control*, 2, 329-52.
- Granger, C.W.J. (1988) "Some Recent Developments in a Concept of Causality", *Journal of Econometrics*, 39, 199-211.
- Greenberg, David F. (editor) (1993) **Crime and Capitalism: Readings in Marxist Criminology**, Temple University Press, Filadélfia.
- Greene, William H. (1997) **Econometric Analysis**, Prentice Hall, Nova Jersey.

- Griffith, Daniel A. (1987) **Spatial Autocorrelation: A Primer**, Resource Publications in Geography, Washington.
- Griffith, Daniel A. (1988) **Advanced Spatial Statistics: Special Topics in the Exploration of Quantitative Spatial Data Series**, Kluwer Academic Publishers, Dordrecht.
- Griffith, Daniel A. (1992) "A Spatially adjusted N-way ANOVA Model", *Regional Science and Urban Economics*, vol.22, nº 3, setembro.
- Griffith, Daniel A. (1995) "The General Linear Model and Spatial Autoregressive Models", in: L. Anselin e R.J.G.M. Florax (editores) **New Directions in Spatial Econometrics**, Springer, Berlim, 273-300.
- Hamilton, James D. (1994) **Time Series Analysis**, Princeton University Press, Princeton.
- Hellman, Daryl A. e Neil O. Alper (1997) **Economics of Crime: Theory and Practice**, 4ª ed., Simon & Schuster, Needham Heights.
- Hicks, John (1979) **Causality in Economics**, Basic Books, Nova York.
- Hoenack, Stephen A. e William C. Weiler (1980) "A Structural Model of Murder Behavior and the Criminal Justice System", *American Economic Review*, vol.70, nº 3, junho, 327-41.
- Holahan, William L. (1998) "Getting Tough on Crime: Exercises in Unusual Difference Curves", *Journal of Economic Education*, inverno, 14-22.
- Horton, Paul B. e Gerald R. Leslie (1970) **The Sociology of Social Problems**, 4ª ed., Appleton-Century-Crofts, Nova York.
- Hotelling, Harold (1929) "Stability in Competition", *The Economic Journal*.
- Hume, David (1777) **An Inquiry Concerning Human Understanding**, editado por Bobbs-Merril, 1995.
- Hsiao, Cheng (1990) **Analysis of Panel Data**, Cambridge University Press, Cambridge.
- Hsieh, Ching-Yao e Meng-Hua Ye (1991) **Economics, Philosophy and Physics**, Sharpe, Armonk.
- Isard, Walter (1956) **Location and Space-Economy**, The M.I.T. Press, Cambridge.
- Jones, R.H. (1993) **Longitudinal Data with Serial Correlation: A State-Space Approach**, Chapman & Hall, Londres.
- Kant, Immanuel (1781) **Critique of Pure Reason**, tradução de J.M.D. Meiklejohn, editado por Collier & son, Nova York, 1966.

- Kelejian, Harry H. e Dennis P. Robinson (1992) "Spatial Autocorrelation: A New Computationally Simple Test With an Application to Per Capita County Police Expenditures", *Regional Science and Urban Economics*, vol.22, nº 3, setembro.
- Kelejian, Harry H. e Dennis P. Robinson (1995) "Spatial Correlation: A Suggested Alternative to the Autoregressive Model", in: L.Anselin e R.J.G.M. Florax (editores) *New Directions in Spatial Econometrics*, Springer, Berlim, 75-95.
- Kern, Stephen (1983) *The Culture of Time and Space 1880-1918*, Harvard University Press, Cambridge.
- Kuhn, Thomas S.(1995) *A Estrutura das Revoluções Científicas*, Perspectiva, São Paulo.
- LeSage, James P. (1995) "A Multiprocess Mixture Model to Estimate Space-Time Dimensions of Weekly Pricing of Certificates of Deposits", in: L.Anselin e R.J.G.M. Florax (editores) *New Directions in Spatial Econometrics*, Springer, Berlim, 359-88.
- Levitt, Steven (1996) "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Legislation", *Quarterly Journal of Economics*, 111, 319-51.
- Levitt, Steven (1997) "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime", *American Economic Review*, vol. 87, nº3, junho, 270-90.
- Lisboa, Marcos B. e Mônica V. Andrade (2000) "Desesperança de Vida: Homicídio em Minas Gerais, Rio de Janeiro e São Paulo: 1981 a 1997", mimeo, programa de seminários acadêmicos IPE-USP, seminário 08/2000.
- Macedo, Paulo B.R. (1998) "Hedonic Price Models with Spatial Effects: an Application to the Housing Market of Belo Horizonte, Brazil", *Revista Brasileira de Economia*, 52(1), 63-81.
- McDonald, John F. (1997) *Fundamentals of Urban Economics*, Prentice Hall, Nova Jersey.
- McMillen, Daniel P. (1995) "Spatial Effects in Probit Models: A Monte Carlo Investigation", in: L.Anselin e R.J.G.M. Florax (editores) *New Directions in Spatial Econometrics*, Springer, Berlim, 189-228.
- McMillen, Daniel P. e John F. McDonald (1991) "A Markov Chain Model for Zoning Change", *Journal of Urban Economics*, 30, 257-70.
- Neher, Philip A. (1978) "The Pure Theory of the Muggery", *American Economic Review*, vol.68, nº 3, junho, 437-45.
- Paelinck, Jean e L.Klaassen (1979) *Spatial Econometrics*, Saxon House, Farnborough.
- Pastore, José, Denise F. Rocca e Liliana Pezzin (1991) *Crime e Violência Urbana*, IPE-FIPE-USP, São Paulo.

- Pezzin, Liliana (1986) **Criminalidade Urbana e Crise Econômica: O Caso de São Paulo**, IPE-USP, São Paulo.
- Pfeifer, Phillip E. e Stuart J. Deutsch (1980a) "A Three-Stage Iterative Procedure for Space-Time Modeling", *Technometrics*, vol.22, nº 1, fevereiro, 35-47.
- Pfeifer, Phillip E. e Stuart J. Deutsch (1980b) "Identification and Interpretation of First Order Space-Time ARMA Models", *Technometrics*, vol.22, nº 3, agosto, 397-408.
- Phillips, Llad (1973) "Crime Control: the Case for Deterrence" in: **The Economics of Crime and Punishment**, American Enterprise Institute, Washington, 65-84.
- Pindyck, Robert S. e Daniel L. Rubinfeld (1998) **Econometric Models and Economic Forecasts**, Irwin McGraw-Hill, Boston.
- Polinsky, A. Mitchell e Steven Shavell (1979) "The Optimal Tradeoff Between the Probability and Magnitude of Fines", *American Economic Review*, vol.69, nº 5, dezembro, 880-91.
- Ponsard, Claude (1954) **Économie et Espace: Essai d'Integration du Facteur Spatial dans l'Analyse Économique**, Sedes, Paris.
- Rubin, Paul H. (1997) "The Economics of Crime" in: N.O. Alper. e D.A. Hellman (editores) **Economics of Crime: A Reader**, 2ª ed., Simon & Schuster, Needham Heights, 1-15.
- Sah, Raaj (1991) "Social Osmosis and Patterns of Crime", *Journal of Political Economy*, 99, 1272-95.
- Simon, J.L. (1970) "The concept of causality in Economics", *Kyklos*, 23.
- Student (1907) "On the error of counting with a haemocytometer", *Biometrika*, 5, 351-60.
- Tittle, Charles R. (1973) "Punishment and Deterrence of Deviance", in: **The Economics of Crime and Punishment**, American Enterprise Institute, Washington, 85-102.
- Wilson, A.G.(1970) **Entropy in Urban and Regional Modelling**, Pion, Londres
- Witte. Ann D. (1980) "Estimating the Economic Model of Crime With Individual Data", *Quarterly Journal of Economics*, 94, 57-8.