

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

Maria Viviana de Freitas Cabral

Avaliação do impacto do INFOCRIM sobre as taxas de homicídios dos municípios paulistas: uma aplicação do método de diferenças em diferenças espacial

**Juiz de Fora
2016**

Maria Viviana de Freitas Cabral

Avaliação do impacto do INFOCRIM sobre as taxas de homicídios dos municípios paulistas: uma aplicação do método de diferenças em diferenças espacial

Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Juiz de Fora - PPGE/UFJF - como requisito parcial para a obtenção do título de Doutora em Economia.

Orientador: Prof. Dr. Eduardo Simões de Almeida

**Juiz de Fora
2016**

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Cabral, Maria Viviana de Freitas.

Avaliação do impacto do INFOCRIM sobre as taxas de homicídios dos municípios paulistas : uma aplicação do método de diferenças em diferenças espacial / Maria Viviana de Freitas Cabral. - 2016.

120 f. : il.

Orientadora: Eduardo Simões de Almeida

Tese (doutorado) - Universidade Federal de Juiz de Fora, Faculdade de Economia. Programa de Pós-Graduação em Economia, 2016.

1. Economia do crime. 2. Avaliação de política de segurança pública. 3. Transbordamentos espaciais. 4. Diferenças-em diferenças espacial. I. Almeida, Eduardo Simões de, orient. II. Título.

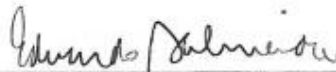
Maria Viviana de Freitas Cabral

Avaliação do impacto do INFOCRIM sobre as taxas de homicídios dos municípios paulistas: uma aplicação do método de diferenças em diferenças espacial

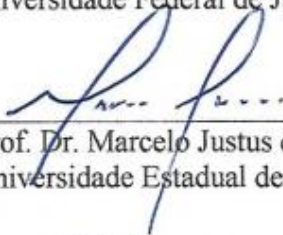
Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Juiz de Fora - PPGE/UFJF - como requisito parcial para a obtenção do título de Doutora em Economia.

Aprovada em 07/06/2016

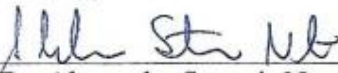
BANCA EXAMINADORA



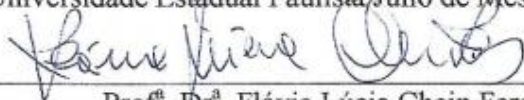
Prof. Dr. Eduardo Simões de Almeida (Orientador)
Universidade Federal de Juiz de Fora



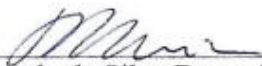
Prof. Dr. Marcelo Justus dos Santos
Universidade Estadual de Campinas



Prof. Dr. Alexandre Sartoris Neto
Universidade Estadual Paulista Júlio de Mesquita Filho



Prof.ª Dr.ª Flávia Lúcia Chein Feres
Universidade Federal de Juiz de Fora



Prof. Dr. Ricardo da Silva Freguglia
Universidade Federal de Juiz de Fora

DEDICATÓRIA

Ao meu marido Joilson e ao fruto do nosso amor, Mariana. Vocês são a família que Deus me permitiu escolher, o motivo e a inspiração para a construção diária da nossa felicidade. Amo vocês cada dia mais!!!

AGRADECIMENTOS

A Deus, em primeiro lugar, por estar presente em todos os momentos da minha vida e à Nossa Senhora por me proteger.

Ao meu querido orientador, Prof. Dr. Eduardo Almeida, por todo suporte, dedicação e confiança na condução deste trabalho.

À Prof^a. Dra. Flávia Lúcia Chein Feres, ao Prof. Dr. Marcelo Justus dos Santos e ao Prof. Dr. Ricardo da Silva Freguglia pelas críticas e sugestões na versão de qualificação que muito enriqueceram este trabalho. Ao Prof. Dr. Alexandre Sartoris Neto que, juntamente com os demais professores, engrandeceu a banca examinadora com sua presença.

A todos os professores do Programa de Pós-Graduação em Economia (PPGE/FE/UFJF) pela expressiva contribuição para minha formação acadêmica. A todos os funcionários da secretaria do PPGE pela gentileza e dedicação.

Aos amigos do doutorado, Juliana, André e Hilton pela agradável convivência ao longo desses anos nos quais prevaleceu a amizade, a descontração, o companheirismo.

Às minhas amigas, Eliane (Lili), Maraísa, Maísa, Vanessa e Carolina que, mesmo longe, acompanharam o processo de concretização deste trabalho, dando todo carinho e incentivos necessários.

Aos colegas de trabalho do DeCE/ICSA/UFRRJ. Em especial, Melk, Alexandre e Maxwell pela amizade e à Aninha e Everlam pela disponibilidade em ajudar.

À minha família: meus pais, Milton e Teresinha, por todo amor e dedicação; minha querida irmã Lili pela amizade sincera; meus irmãos, Cristiano, Fabiano e Liciano pelo apoio; aos meus sobrinhos Arthur e Diogo pelo amor despretenso; à minha cunhada Lyse Anne pela torcida. À minha “mãe carioca” Ana pelo carinho e amizade. Obrigada família pelo amor, incentivo e apoio incondicionais. Amo vocês!

Ao meu marido Joilson. Meu companheiro e meu amigo. Meu colega de trabalho e parceiro profissional. Minha melhor companhia sempre. Você foi fundamental para a concretização desse trabalho. À nossa filha Mariana, nosso maior motivo para a constante busca do aprimoramento pessoal e profissional. Agradeço a Deus pela nossa família: vocês me ensinaram o que é o amor no sentido mais sincero da palavra. Essa conquista é nossa!

RESUMO

A reversão dos índices de crime no estado de São Paulo tem despertado o interesse dos pesquisadores em verificar as possíveis causas para a redução de 67,1% nas taxas de homicídios no período de 2000 a 2010. Diversos fatores têm sido apontados para explicar esse fenômeno, porém não foi encontrado na literatura, até o momento, estudos que avaliassem o Sistema de Informações Criminais – INFOCRIM, adotado por 67 municípios paulistas até o ano de 2010. Desse modo, este trabalho tem o objetivo de testar a hipótese de que o INFOCRIM tenha contribuído para a redução das taxas de homicídios entre 2000 e 2010, controlando por diversos determinantes do crime e por outras iniciativas de segurança pública. Para isso, foi realizada uma avaliação de tratamento por meio da abordagem das diferenças-em-diferenças espacial (SDID), método pioneiro capaz de modelar a migração do crime das regiões tratadas para regiões vizinhas. O grupo de tratamento foi definido como os 67 municípios contemplados pelo INFOCRIM, ao passo que o grupo de controle foi definido como os 548 municípios paulistas remanescentes na amostra. Por meio de um painel de dados de dois períodos (2000 e 2010), o modelo de defasagem espacial (SDID-SAR) evidenciou a existência de transbordamentos espaciais (migração de crime) entre os municípios paulistas e um efeito redutor do crime devido ao INFOCRIM. Em termos de efeitos totais, o INFOCRIM reduziu a taxa de homicídios em 6,183 entre 2000 e 2010. Quanto aos efeitos diretos, o Programa reduziu a taxa de homicídios em 3,745, enquanto, em termos de efeitos indiretos, o INFOCRIM diminuiu a taxa de homicídios em 2,437 durante esse período. Isso implica que a interação espacial existente entre os agentes reforçou o efeito médio do INFOCRIM. Os resultados corroboraram a relação positiva entre crime e desemprego e a relação inversa entre crime e educação. A melhoria da equidade social contribuiu para a redução da criminalidade violenta enquanto o adensamento populacional apresentou sinal contrário ao esperado. Há um efeito dissuasor indireto sobre as taxas de homicídios decorrente da existência de guarda municipal armada. Em termos de vidas poupadas, o Estado de São Paulo foi capaz de minimizar os prejuízos sociais decorrentes da criminalidade, uma vez que a implementação do INFOCRIM evitou 2.546 homicídios de 2000 até 2010. Portanto, o INFOCRIM pode ser considerado um Programa exitoso no combate à criminalidade letal.

PALAVRAS-CHAVE: Economia do crime. Avaliação de política de segurança pública. Transbordamentos espaciais. Diferenças-em-diferenças espacial.

CLASSIFICAÇÃO JEL: C21, I38, K42.

ABSTRACT

The reversal of the trend of crime rates in the state of São Paulo has arisen the interest of researchers to verify the possible causes for the reduction of 67.1% in homicide rate from 2000 to 2010. Several factors have been suggested to explain this phenomenon, but it was not found in the literature, so far, researches to evaluate the Criminal Information System – INFOCRIM, adopted by 67 municipalities in São Paulo until 2010. Thus this work is aimed at testing the hypothesis that INFOCRIM has contributed to the reduction in homicide rate between 2000 and 2010, controlling for determinants of crime and other public security actions. To do so, a treatment evaluation is performed using the spatial difference-in-differences (SDID) approach, a pioneering method able to model the crime migration from the treated regions to their neighboring regions. The treatment group is defined as being the 67 municipalities with INFOCRIM whereas the control group is defined as being 548 municipalities remaining in the sample. Using two-period panel data (2000 and 2010), the spatial lag model (SDID-SAR) indicated the existence of spatial spillovers (crime migration) among municipalities in the São Paulo state and a crime reducing effect of INFOCRIM. As to total effects, INFOCRIM reduced homicide rate by 6.183 over the period 2000-2010. In terms of direct terms, the program decreased homicide rate by 3.745, while, in terms of indirect effects, INFOCRIM reduced homicide rate by 2.437 during this period. . This implies that the spatial interaction between agents enhanced the average effect of INFOCRIM. These findings confirmed the positive relation between crime and unemployment and the inverse relation between crime and education. The improvement of social equity contributed to the reduction of violent crime while the population density showed a sign contrary to theoretically expected. There is an indirect effect on homicide rate due to the existence of municipal guard. In terms of lives saved, the Government of the State of São Paulo was able to minimize the social loss from offenses, since the implementation of INFOCRIM has avoided 2,546 homicides over the period 2000-2010. Therefore, the INFOCRIM can be considered successful in fighting lethal crime.

KEYWORDS: Economic theory of crime. Public security policy evaluation. Spatial spillovers. Spatial difference-in-differences.

LISTA DE ILUSTRAÇÕES

Figura 1.1 – Evolução da taxa de homicídios por cem mil habitantes no Brasil e no estado de São Paulo, 1990 a 2010.	11
Figura 3.1 - Resumo do procedimento de estimação dos modelos SDID	63
Figura 3.2 - Grupos de tratamento e de controle para a implementação do DID e SDID.	69
Figura 3.3 - Tendência das taxas de homicídios por cem mil habitantes dos grupos de tratamento e de controle, 2000 a 2010.....	70
Figura 3.4 – Mapas para a taxa de homicídios por cem mil habitantes, em nível (a e b), da criminalidade nos municípios do Estado de São Paulo, 2000 e 2010.	86
Figura 3.5 - Mapas de <i>clusters</i> e de significância para a taxa de criminalidade por cem mil habitantes nos municípios do Estado de São Paulo, 2000 e 2010.	90

LISTA DE QUADROS E TABELAS

Quadro 2.1 - Resumo dos trabalhos empíricos sobre avaliação de políticas de segurança pública.	25
Quadro 3.1 – Resumo esquemático do método DID.	37
Quadro 3.2- Resumo dos principais modelos de diferenças em diferenças espaciais.....	62
Quadro 3.3 - Resumo das variáveis utilizadas no modelo empírico.	78
Tabela 3.1 - Estatística descritiva dos dados no período <i>ex ante</i> ao tratamento, 2000.....	80
Tabela 3.2 - Matriz de correlação das variáveis do modelo empírico.	81
Tabela 3.3 – Estatísticas <i>I</i> de Moran uni e bivariada, 2000 e 2010.	88
Tabela 3.4 – Teste <i>join count</i> para a autocorrelação espacial.	89
Tabela 4.1 – Resultado das estimações dos modelos de diferenças em diferenças, DID.....	93
Tabela 4.2 – Resultados do teste de autocorrelação espacial <i>I</i> de Moran para painel nos resíduos do modelo DID estimado por efeitos fixos.....	120
Tabela 4.3 – Resultados das estimações dos modelos de diferenças em diferenças espacial, SDID.....	97
Tabela 4.4 – Decomposição do efeito total do modelo SDID-SAR em efeitos direto e indireto.	99

SUMÁRIO

	Páginas
1 INTRODUÇÃO	10
2 REVISÃO DA LITERATURA EMPÍRICA.....	15
2.1 EVIDÊNCIAS EMPÍRICAS DO IMPACTO DAS POLÍTICAS DE SEGURANÇA PÚBLICA.....	15
3 METODOLOGIA	31
3.1 DIFERENÇAS EM DIFERENÇAS.....	33
3.2 DIFERENÇAS EM DIFERENÇAS ESPACIAL.....	38
3.2.1 Modelo SDID-SAR	43
3.2.2 Modelo SDID-SEM.....	48
3.2.3 Modelo SDID-SAC	52
3.2.4 Modelo SDID-SLX	55
3.2.5 Modelo SDID-SDM	57
3.2.6 Modelo SDID-SDEM.....	59
3.2.7 Resumo dos métodos de estimação dos modelos SDID.....	61
3.3 A TEORIA ECONÔMICA DA ESCOLHA RACIONAL: DA OFERTA INDIVIDUAL À OFERTA AGREGADA DO CRIME	64
3.4 ESTRATÉGIA EMPÍRICA E DESCRIÇÃO DAS VARIÁVEIS.....	68
3.5 A CRIMINALIDADE EVITADA PELO INFOCRIM.....	83
3.6 ANÁLISE EXPLORATÓRIA DOS DADOS ESPACIAIS	85
4 ANÁLISE DO IMPACTO DO INFOCRIM SOBRE A CRIMINALIDADE .	92
5 CONSIDERAÇÕES FINAIS E IMPLICAÇÕES POLÍTICAS	103
REFERÊNCIAS	107

1 INTRODUÇÃO

A criminalidade é um dos problemas mais complexos enfrentados pela sociedade moderna por se tratar de um problema multidimensional de natureza social, política e econômica. É um problema social por afetar diretamente a qualidade e expectativa de vida e um problema político, pois as ações preventivas e de combate exigem a participação ativa dos governos e a alocação dos recursos públicos escassos. A criminalidade consiste também em um problema econômico porque, de um lado, a sua intensidade está associada às condições econômicas e, de outro, limita o potencial de desenvolvimento econômico por deteriorar o capital humano (BEATO FILHO, 1999; ARAÚJO JR e FAJNZYLBER, 2001).

A investigação econômica do crime tem sua origem em Beccaria (1764) e Bentham (1789). Na década de 1960, observa-se um crescimento do debate sobre a criminalidade do ponto de vista econômico com os estudos de Fleisher (1963, 1966), Ehrlich (1967) e outros. Entretanto, a formalização dos modelos econômicos do crime é creditada a Becker (1968) e Ehrlich (1973) sob a premissa de que os criminosos se comportam racionalmente no sentido de maximizar o seu bem-estar econômico. A partir desses trabalhos seminais, diversos estudos têm sido empreendidos para uma melhor compreensão dos determinantes do crime (EHLICH, 1973, 1996; BLOCK e HEINEKE, 1975; FREEMAN, 1983, 1996; KELLY, 2000; LEVITT, 2004) e dos custos da criminalidade (BOURGUIGNON e MORRISON, 2000; BUVINIC e MORRISON, 2001).

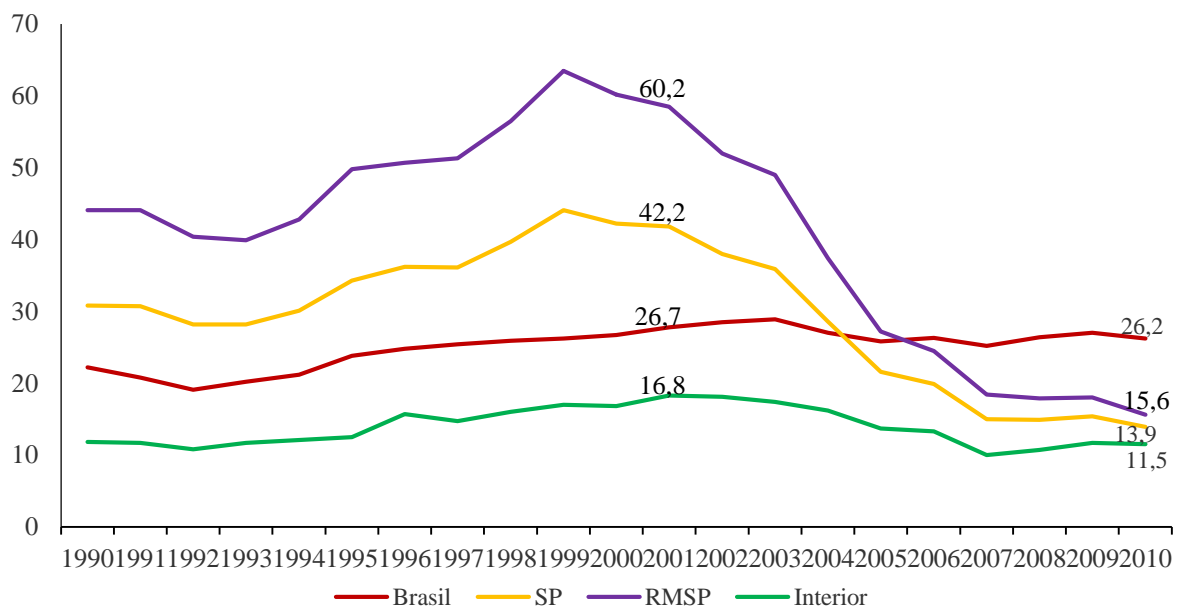
Entretanto, apesar da importância do tema, o Brasil iniciou a primeira década do novo milênio com indicadores de criminalidade preocupantes. Em 2000, o País possuía a quinta maior taxa de homicídios na América Latina, com 26,7 casos para cada cem mil habitantes¹, ficando atrás de Colômbia (66,7), El Salvador (59,8), Honduras (51,1) e Venezuela (32,9). Em 2010, o Brasil permaneceu na quinta posição entre os países mais violentos da América Latina com uma taxa de homicídios de 26,2 para cada cem mil habitantes. Apesar da melhora na taxa de homicídios, o Brasil lidera o *ranking* de homicídios em números absolutos no período com 45.336 registros em 2000 e 49.978 em 2010 (UNODC, 2012). Diante dessas estatísticas, a criminalidade é um dos principais problemas enfrentados pelo Brasil nos anos recentes.

¹ A Organização Mundial da Saúde estabelece que uma taxa acima de 10 homicídios para cada 100 mil habitantes pode ser considerada epidêmica (UNDP, 2013).

Na contramão das taxas de homicídios elevadas e relativamente estáveis do Brasil na primeira década dos anos 2000, a trajetória das estatísticas criminais no estado de São Paulo tem-se mostrado declinante. Entre os anos de 2000 e 2010, as taxas de homicídios por cem mil habitantes no estado diminuíram 67,1%, passando de 42,2 em 2000 para 13,9 casos de homicídios para cada cem mil habitantes em 2010 (Figura 1.1). Em números absolutos, a redução dos homicídios no estado foi de 258%, contabilizando 14.745 casos de homicídios dolosos em 2000 contra 5.713 casos notificados em 2010.

Essa redução sistemática das taxas de homicídios propiciou uma posição de destaque no contexto nacional: de 4º estado mais violento em 2000, o estado de São Paulo passa a ocupar a 25ª posição, tornando-se um dos 3 estados com menores índices de criminalidade letal do País. A reversão da criminalidade letal fez com que a Organização das Nações Unidas reconhecesse o estado como um caso bem-sucedido no cenário mundial, sendo comparado a outros casos de sucesso de redução da criminalidade, como Nova Iorque e Bogotá.

Figura 1.1 – Evolução da taxa de homicídios por cem mil habitantes no Brasil e no estado de São Paulo, 1990 a 2010.



Fonte: Elaboração Própria a partir de dados do SIM/SVS/MS.

Como visualizado por intermédio da Figura 1.1 e observado por Hartung (2009), apesar de a redução nos homicídios em São Paulo ter ocorrido no estado como um todo, a queda mais acentuada foi observada nos municípios maiores. Waiselfisz (2011) revela que as mudanças

mais expressivas ocorreram nas regiões metropolitanas de São Paulo, Campinas, na Baixada Santista e no litoral norte do estado.

Com vistas a explicar o fenômeno da redução abrupta e ininterrupta das taxas de homicídios em São Paulo, Kahn e Zanetic (2005) discutem o papel dos municípios no combate e prevenção à criminalidade e concluem que a divulgação do Disque Denúncia, a implementação da lei de restrições ao consumo de álcool e de projetos de inclusão social, bem como a alocação adequada das guardas municipais podem ser consideradas possíveis explicações para o fenômeno. Para verificar formalmente o impacto de tais iniciativas na esfera municipal, Scripilliti (2006) realizou uma avaliação de tratamento e concluiu que as restrições ao consumo de álcool e a existência de guardas municipais e de Secretaria de Segurança Pública Municipal não contribuíram para a redução de crimes contra o patrimônio e das taxas de homicídios dolosos nos municípios paulistas entre os anos de 1997 e 2003. De acordo com as conclusões do estudo, somente a maior divulgação do Disque Denúncia teria impacto sobre os homicídios dolosos.

Já Mello e Schneider (2010) apontam a transição demográfica ocorrida no estado como fator determinante para a queda da criminalidade, enquanto Biderman *et al.* (2010) mostram que as restrições ao consumo de álcool em algumas cidades de São Paulo causaram uma redução de, aproximadamente, 10% nas taxas de homicídios. Os estudos de Hartung (2009), Cerqueira (2010), Cerqueira e Mello (2012), Justus e Kassouf (2012, 2013) encontraram evidências de que o Estatuto do Desarmamento tenha contribuído para a redução da criminalidade em São Paulo. Para Dias (2011) e Willis (2013, 2015), a redução dos homicídios no estado e na cidade de São Paulo, respectivamente, observada nos últimos anos pode ser creditada à “hipótese PCC”. Tal hipótese apregoa que a monopolização do “mundo do crime” via expansão e consolidação da facção criminosa Primeiro Comando da Capital (PCC) no sistema carcerário paulista foi capaz de desenvolver mecanismos de controle social que produziram uma drástica redução do uso da violência física nos conflitos interpessoais e, assim, possibilitaram a imposição da paz nestes territórios.

Por sua vez, o Governo do estado de São Paulo aponta a implementação da política de segurança pública estadual iniciada nos anos 2000 como o principal fator responsável pela redução do número de homicídios entre 2000 e 2010. De acordo com a Secretaria de Segurança Pública de São Paulo (SSP/SP), a adoção e expansão de sistemas tecnológicos para a investigação policial e para os policiamentos preventivo e ostensivo foi um importante investimento do Governo de São Paulo na segurança pública. Nesse sentido, os estudos qualitativos de Kahn (2007), Nóbrega Junior *et al.* (2009) e Goertzel e Kahn (2009) consideram

a melhoria da gestão estadual de segurança pública como um dos fatores para explicar a trajetória declinante das taxas de homicídios em São Paulo.

Apesar de a literatura apontar diversos fatores para explicar a redução da taxa de homicídios no estado, não foi encontrado, até o momento, nenhum estudo que avaliasse quantitativamente os sistemas tecnológicos adotados pelas polícias do estado de São Paulo a partir dos anos 2000. O sistema tecnológico capaz de reunir todas as informações digitais acerca da criminalidade é o Sistema de Informações Criminais, conhecido como INFOCRIM. O INFOCRIM visa aumentar a eficiência das polícias por meio do sistema de georreferenciamento do crime baseado na informatização do registro de estatísticas de ocorrências criminais, além de promover a avaliação constante do desempenho da polícia.

Criado em 1999, o lançamento de uma rede informatizada no ano de 2000 marca o início do INFOCRIM na cidade de São Paulo. Em 2003, a SSP/SP assinou um convênio com a Prefeitura de São Paulo, viabilizando o acesso ao INFOCRIM na Capital, além de concluir a implantação do Sistema nos 38 municípios da Região Metropolitana (RMSP) e nas cidades de Campinas e Santos. Até o ano de 2010, o Programa estava efetivamente disponível para um total de sessenta e sete municípios do estado de São Paulo² (SSP/SP, 2010).

Diante do exposto, o objetivo desta tese é testar a hipótese de que o INFOCRIM, *proxy* para a política de segurança pública estadual, tenha contribuído para a redução das taxas de homicídios no período de 2000 a 2010. Visando atingir o objetivo proposto, será realizada uma avaliação de tratamento por meio dos métodos de diferenças em diferenças (DID) e diferenças em diferenças espacial (SDID) para um painel de dados de dois períodos (2000 e 2010). Para implementar tais métodos, o grupo de tratamento foi definido como os sessenta e sete municípios onde foi implementado o INFOCRIM e o grupo de controle como os demais municípios entre os 615 municípios paulistas contemplados na amostra.

A partir dessas avaliações de tratamento, é possível inferir se o INFOCRIM implementado em São Paulo afetaria o comportamento racional do potencial criminoso de modo a reduzir a oferta de crime no estado conforme a lógica da teoria econômica do crime de Becker (1968) e Ehrlich (1973). Portanto, esta tese pretende contribuir para a literatura de economia do crime ao avaliar o impacto do INFOCRIM sobre as taxas de homicídios nos municípios de São Paulo por meio de um método de avaliação de tratamento pioneiro (SDID). O método SDID flexibiliza algumas hipóteses da metodologia DID, difíceis de serem garantidas no caso de avaliação de políticas de segurança pública. O SDID é capaz de modelar os possíveis

² A lista dos sessenta e sete municípios que implementaram o INFOCRIM com a respectiva data de implantação do Programa encontra-se no Apêndice A.

transbordamentos da criminalidade³ entre as unidades regionais analisadas, o que propiciará a identificação do efeito causal do INFOCRIM sobre a dinâmica da criminalidade no estado de São Paulo.

Além dessa introdução, a presente tese está disposta da seguinte maneira: no capítulo 2 é realizada uma revisão empírica sobre a avaliação de diferentes tipos de políticas de segurança pública nas literaturas internacional e nacional. No capítulo 3 é detalhada a metodologia empregada. Neste terceiro capítulo também encontra-se o modelo econômico do crime para compreender os diversos determinantes da criminalidade, mais especificamente, o impacto de uma política de segurança pública na oferta do crime com vistas a dar suporte teórico à estratégia empírica adotada com a subjacente descrição da base de dados utilizada. O capítulo 4 se atém à análise do impacto do INFOCRIM sobre a criminalidade letal nos municípios de São Paulo. Por fim, o capítulo 5 apresenta as considerações finais e as implicações políticas da tese.

³ Na literatura de economia do crime, esses transbordamentos espaciais são também denominados de migrações de crime ou deslocamentos de crime.

2 REVISÃO DA LITERATURA EMPÍRICA

Como o objetivo central da tese é analisar o impacto do INFOCRIM sobre a criminalidade violenta nos municípios paulistas, torna-se importante conhecer como a avaliação do impacto de políticas de segurança pública é tratada tanto na literatura internacional quanto na literatura nacional, considerando cada um dos diferentes tipos de políticas. Para tanto, a seção 2.1 do presente capítulo apresenta os estudos empíricos que avaliaram diferentes tipos de política de segurança pública com o intuito de explicitar a contribuição do presente trabalho à literatura vigente.

2.1 EVIDÊNCIAS EMPÍRICAS DO IMPACTO DAS POLÍTICAS DE SEGURANÇA PÚBLICA

Considerando o impacto da criminalidade sobre o bem-estar econômico e social, a avaliação de programas tem despertado o interesse dos pesquisadores nos anos recentes. Esse interesse é motivado pelo fato de que a avaliação de programas é um mecanismo de melhoria do processo de tomada de decisões, visto que a avaliação permite ao governante fazer inferências acerca dos resultados de um programa específico, informação que pode ser utilizada para melhorar a concepção ou implementação do programa, para fundamentar decisões e para melhorar a prestação de contas sobre políticas públicas.

De forma geral, boa parte das políticas de segurança pública executadas em diversos países, principalmente países latino-americanos, tem inspiração nos elementos-chave das estratégias políticas norte-americanas. Nos Estados Unidos, expressivas mudanças nas políticas de combate e prevenção à criminalidade ocorreram nos últimos trinta anos em decorrência de novas tecnologias, novas perspectivas sobre policiamento e recorrentes avaliações dos resultados das políticas de segurança pública. Santos (2013) descreve seis diferentes estratégias de políticas de segurança pública, quais sejam: o modelo padrão de policiamento; policiamento comunitário; policiamento *Broken Windows*; policiamento *hot spots*; COMPSTAT e policiamento orientado para os problemas.

O elemento central do modelo padrão de policiamento envolve a aplicação rigorosa da lei utilizando, principalmente, o policiamento ostensivo. O reduzido número de trabalhos que

avaliaram esse tipo de política mostraram reduzido impacto para o combate da criminalidade (SHERMAN *et al.*, 1997; SKOGAN e FRYDL, 2004; WEISBURD e ECK, 2004).

Quanto ao policiamento comunitário (SKOLNICK e BAYLEY, 2002), uma das políticas mais adotadas, a ideia básica é a parceria entre a polícia e a comunidade para tratar de forma pró-ativa as condições que originam os problemas de segurança pública. De maneira geral, os resultados dos estudos que avaliaram tais políticas nos Estados Unidos mostraram que o policiamento comunitário parece não surtir efeito sobre a prevenção da criminalidade. No entanto, demonstrou-se pertinente para reduzir a sensação de medo da criminalidade (WEISBURD e ECK, 2004).

No contexto do policiamento comunitário estão inseridos as Unidades de Polícia Pacificadora (UPPs) no Rio de Janeiro e o Programa Ronda do Quarteirão em Fortaleza. As UPPs integram um projeto da secretaria estadual de segurança pública do Estado do Rio de Janeiro, criado em 2008, com o objetivo de instituir polícias comunitárias em favelas como meio de recuperar territórios ocupados por grupos criminosos nessas localidades e pacificar as áreas, sendo sua estratégia fundamentada na parceria entre a população e as instituições de segurança pública (CARDOSO *et al.*, 2016). Para avaliar o impacto das UPPs sobre a criminalidade no estado do Rio de Janeiro no período de 2005 a 2012, Ferreira (2013) implementou a metodologia de diferenças em diferenças definindo o grupo de controle como todos os departamentos de polícia (DPs) do interior do estado e como grupos de tratamento os DPs da Baixada Fluminense e da Capital. De acordo com os resultados do estudo, os homicídios diminuíram tanto na Baixada quanto na Capital, sendo a queda mais acentuada verificada nos DPs localizados na Capital. Contudo, o Programa não alterou os índices de latrocínios cometidos no estado tampouco reduziu os índices de roubo a veículos na região da Baixada Fluminense.

Já o Programa “Ronda do Quarteirão”, implementado em novembro de 2007 nos bairros da Região Metropolitana de Fortaleza, capital do estado do Ceará, objetiva prevenir e coibir o cometimento de atos ilícitos. Para verificar se tal objetivo foi atingido, Suliano e Oliveira (2013) analisaram o impacto do aumento do efetivo policial a partir da implantação do Programa “Ronda do Quarteirão” no período compreendido entre os meses de janeiro-maio de 2007 e janeiro-maio de 2008 por intermédio da metodologia de diferenças em diferenças. Os resultados sugerem que o aumento recente do efetivo policial nos bairros tratados apresentou uma redução de 57% nas taxas de roubos. Sendo assim, os autores sustentam a hipótese do chamado efeito dissuasão ao ratificar que o aumento do efetivo policial alterou os incentivos dos criminosos, inibindo a oferta de crimes por tais indivíduos.

Apesar de a atribuição principal da guarda municipal ser a proteção do patrimônio público municipal⁴, essa iniciativa municipal também pode ser inserida no contexto de policiamento comunitário, uma vez que atuação da guarda municipal pode ter efeito indireto sobre a prevenção da criminalidade. Para verificar o impacto da existência de guardas municipais sobre a criminalidade, Scripilliti (2006) analisou os crimes contra patrimônio e homicídios dolosos no período compreendido entre os anos de 1997 e 2003 nos municípios paulistas. Por intermédio de um pseudo-painel de dois períodos, o autor utilizou a metodologia de diferenças em diferenças, definindo o grupo de tratamento como os municípios onde as guardas municipais foram implementadas e como grupo de controle os demais municípios paulistas. As conclusões do estudo sugerem que a existência de guardas municipais não contribuiu para a redução dos crimes contra o patrimônio, bem como para a redução do número de homicídios dolosos. Scripilliti (2006) também analisou outras iniciativas na esfera municipal como a divulgação do Disque-Denúncia, a Lei Seca e a existência de Secretaria de Segurança Pública Municipal e, de forma geral, tais iniciativas municipais não contribuíram para a redução dos crimes contra o patrimônio. Somente o Disque Denúncia impactaria as taxas de homicídios dolosos no montante de -0,394% a cada denúncia adicional, considerando o nível de significância de 10%.

A teoria *Broken Windows*, desenvolvida por Wilson e Kelling (1982), conhecida também como Tolerância Zero, buscava demonstrar a relação de causalidade entre desordem e criminalidade. De acordo com os princípios da teoria das “Janelas Quebradas”, todo e qualquer ato criminoso deveria ser veementemente reprimido a fim de evitar a ocorrência de delitos de maior potencial ofensivo: policiar incivildades (ou microviolências) e manter a ordem. O termo “janelas quebradas” é uma metáfora que alude ao fato de que, se uma janela quebrada é deixada sem conserto, induziria as pessoas a acreditarem que ninguém se importa, o que levaria os vândalos a depredarem mais janelas, e, eventualmente, praticarem um comportamento criminoso mais sério (SOUSA e KELLING, 2006).

Com o propósito de estender teórica e empiricamente Wilson e Kelling (1982), Skogan (1990) investigou como incivildades influenciaram o crime no nível de bairro. O autor mostrou que a distribuição da desordem em quarenta áreas urbanas dos EUA estava positivamente correlacionada com o nível de pobreza, a composição racial e o grau de instabilidade nestes locais, entre os anos de 1977 e 1983. Somente para um bairro de Newark, Nova Jersey, foram encontradas evidências de que a aplicação rigorosa da lei tenha reduzido a criminalidade.

⁴ A Constituição Federal de 1988 em seu artigo 144, parágrafo 8º dispõe que “os municípios poderão constituir guardas municipais destinadas à proteção de seus bens, serviços e instalações, conforme dispuser a lei”.

Skogan (1992), por sua vez, verificou que a desordem social havia aumentado, apesar dos projetos de segurança pública de repressão rigorosa de quaisquer delitos, em sete cidades norte-americanas.

No início de sua gestão em 1994, o prefeito de Nova Iorque, Rudolph Giuliani, adotou o programa “Tolerância Zero”, passando a exercer repressão sobre qualquer tipo de ato ilícito, inclusive os menos graves. Com a redução das taxas de criminalidade nos anos posteriores, a política de segurança pública adotada foi considerada exitosa por seus *policy makers*. Kelling e Sousa (2001) empregaram um modelo hierárquico linear com dados socioeconômicos e demográficos, de policiamento “*broken windows*” (prisões por contravenção) e uso de drogas no período de 1989 a 1998 em Nova Iorque. De acordo com os autores, o maior impacto sobre os índices de crimes violentos decorreu do policiamento “*broken windows*”, sugerindo o êxito da política de segurança pública implementada.

Nessa mesma perspectiva, Corman e Mocan (2005), por meio de métodos de séries temporais com dados mensais de 1974 a 1999, concluíram que a contribuição de medidas de dissuasão (número de detenções e força policial) é maior do que das variáveis econômicas (taxa de desemprego e salário mínimo real) para a queda da criminalidade na cidade de Nova Iorque. Todavia, as táticas repressivas de policiamento (“Tolerância Zero”) são negativamente associadas com furtos e roubos de veículos, mas não tem impacto significativo sobre outros crimes.

Por outro lado, Donohue e Levitt (2001) chegaram à conclusão de que a redução da criminalidade observada na cidade de Nova Iorque não decorreu da política de segurança adotada, mas em virtude da liberalização do aborto nos Estados Unidos em 1973, dado que milhares de crianças que nasceriam sob condições que poderiam induzi-las à atividade criminosa deixaram de nascer. Os autores concluíram que mais de 50% da redução da taxa de criminalidade deveu-se à legalização do aborto.

Além da legalização do aborto, Levitt (2004)⁵ apontou o aumento do contingente policial e das prisões como fatores para a redução de crimes verificada. Corroborando as conclusões de Bowling (1999) e Blumstein (1995), o autor também sugeriu a redução da “epidemia” de *crack* como um importante fator para a redução da criminalidade nova-iorquina. Ainda, Levitt (2004) atentou para o fato de que a queda das taxas de criminalidade foi observada

⁵ Nesse estudo, seis fatores foram apontados como tendo pouco ou nenhum papel sobre a queda verificada. São eles: o crescimento econômico da década, as mudanças demográficas, melhores estratégias policiais, leis de controle e de porte de armas de fogo e lei da pena de morte.

nos EUA como um todo e não apenas na cidade de Nova Iorque, fato em consonância com o exposto por Eck e Maguire (2000).

Quanto à política de policiamento *hot spots*, a ideia-chave é a de que um número desproporcional de crimes acontece em determinadas áreas em uma cidade conhecidas como “áreas quentes”. Para tanto, a estratégia de combate ao crime consiste em identificar sistematicamente os *hot spots* de modo a empregar respostas ofensivas em tais áreas, como aumento do efetivo policial e de prisões nessas localidades. Entretanto, cabe a ressalva de que a avaliação da política *hot spots* precisa considerar se a mesma realmente controla o crime ou se há apenas o deslocamento do crime para outra área (SANTOS, 2013).

Em resumo, os resultados das diversas avaliações acerca do tipo de política de segurança *hot spots* mostram que a mesma contribui para reduzir crime e desordem (KATZ *et al.*, 2001; WEISBURD e ECK, 2004; SABATES e FEINSTEIN, 2008, BRAGA 2007; BRAGA e BOND, 2008; MACHIN e MARIE, 2011; RATCLIFFE *et al.*, 2011). Os estudos observaram que, mesmo quando um *hot spot* é movido, em outros termos, mesmo quando há transbordamento da criminalidade para outras áreas, esse tipo de política promove alguma redução da criminalidade (SANTOS, 2013).

No contexto nacional, uma das principais características do programa de segurança pública implementado em Pernambuco, o Pacto pela Vida, consiste no estabelecimento de regiões prioritárias de investimento e atuação definidas com base nas taxas de homicídio mais elevadas ou, simplesmente, a prioridade do Programa são os *hot spots*. O estudo de Silveira Neto *et al.* (2013) buscou obter uma estimativa do impacto do programa Pacto pela Vida sobre a taxa de homicídio do Estado entre 2007 e 2011. Para tanto, os autores empregaram a estratégia de construção de um controle sintético, estendida por Abadie *et al.* (2010), para obter um contrafactual da taxa de homicídio de Pernambuco, considerando-se um grupo de 23 estados do Brasil. Os resultados obtidos mostraram que, em média, houve uma redução anual de 9,1 pontos na taxa de homicídio de Pernambuco, quando comparado ao seu controle sintético para o período 2007-2011. Isso equivale a uma redução de, aproximadamente, 17,3% em relação aos níveis criminalidade em 2007 ou cerca de 2.213 crimes evitados em decorrência do Programa.

“COMPUter-STATistics meetings” (BRATTON, 1996) ou “COMParE STATs” (SILVERMAN, 1999), mais conhecido por sua sigla COMPSTAT, é a política de segurança pública específica implementada na cidade de Nova Iorque na década de 1990 e que, nos anos posteriores, foi amplamente adotada e difundida em todos os Estados Unidos. O COMPSTAT baseia-se no uso cartográfico de estatísticas de delinquência, prisões e dados de qualidade de vida de modo a utilizar essas informações para definição de estratégias de prevenção e redução

do crime. Conforme apontado por Willis *et al.* (2007), a maciça adesão deste tipo de programa nos EUA pode ser devida ao fato de parecer uma política bem-sucedida na redução dos índices de criminalidade, afinal, os primeiros três anos de implementação do COMPSTAT em Nova Iorque coincidiu com declínios expressivos na taxa de criminalidade.

Apesar dos indícios de que a estratégia COMPSTAT tenha surtido efeito sobre a redução da criminalidade, Jang *et al.* (2010) salienta que poucas pesquisas empíricas têm avaliado o impacto do Programa sobre a redução do crime. Alguns estudos que avaliaram o Programa COMPSTAT em Nova Iorque encontraram resultados conflitantes sobre a relação causal entre a estratégia COMPSTAT e o declínio nas taxas de criminalidade (KELLING e SOUSA, 2001; ROSENFELD *et al.*, 2005). Kelling e Sousa (2001) encontraram evidências de um aumento significativo em apreensões por contravenção após o início da estratégia COMPSTAT. Por meio de estudo de caso conduzido nas seis maiores delegacias da cidade em 2000, os autores verificaram que os delegados utilizavam a tecnologia COMPSTAT para identificar quando tipos específicos de crime, como furtos e roubos, tornavam-se problemas graves. A partir dessa identificação da gravidade do problema, os delegados promoviam estratégias específicas de combate, resultando na queda da incidência de tais crimes.

Já Rosenfeld *et al.* (2005) compararam três Programas: COMPSTAT em Nova Iorque, Operação *Ceasefire* em Boston e Projeto *Exile* em Richmond na Virginia. Aplicando análise de curvas de crescimento e modelos hierárquicos lineares generalizados para dados das 95 maiores cidades dos EUA no período de 1992 a 2001, os autores observaram que a tendência de homicídios da cidade de Nova Iorque na década de 1990 não diferiu significativamente em relação a outras grandes cidades. De acordo com os resultados do estudo, há evidências de queda mais acentuada nos homicídios para Boston. Ademais, Weisburd e Braga (2006) observaram que o COMPSTAT tem sido implementado concomitantemente com outras iniciativas de prevenção à criminalidade, como tolerância zero ou *hot spots*, o que dificulta a identificação do impacto do COMPSTAT sobre o crime. Eck e Maguire (2000), por sua vez, alertaram para o fato de que a queda observada nas taxas de criminalidade em Nova Iorque e em outras três cidades norte-americanas começou antes da implementação do COMPSTAT.

O artigo seminal de Goldstein (1979) introduz o conceito de policiamento orientado para os problemas cuja ideia é a de que a polícia assuma um papel proativo na identificação, compreensão e resposta aos problemas em suas comunidades. O argumento de Goldstein (1979) sugere que se a polícia tivesse um “foco orientado para os problemas” poderia ser mais eficaz no combate ao crime, uma vez que a polícia deveria coletar novos dados, desenvolver novos métodos de análise, identificar soluções inovadoras e aplicar medidas para avaliar o sucesso de

seus esforços. As evidências obtidas pelos estudos de Skogan e Frýdl (2004) e Weisburd *et al.* (2008) mostraram que a estratégia parece ser uma das mais promissoras na prevenção e combate ao crime.

Uma das mais conhecidas iniciativas do tipo policiamento orientado, o *Boston Gun Project*, foi uma ação orientada para combater um grave problema: a vitimização de jovens em Boston, cidade que experimentou uma epidemia de homicídios de jovens entre o final de 1980 e início de 1990. No início de 1996, teve início a Operação *Ceasefire* com a aplicação rígida da lei sobre armas de fogo ilícitas que abasteciam os jovens e uma tentativa de gerar um forte elemento dissuasor para a violência entre as gangues (estratégia chamada de *pulling levers*). Com o intuito de avaliar a intervenção *Ceasefire* sobre a vitimização de jovens em Boston, Braga *et al.* (2001) analisaram dados mensais de número de homicídios de jovens no período de janeiro de 1991 a maio de 1998, empregando modelos lineares generalizados por meio de regressão de Poisson e chegaram à conclusão de que a operação *Ceasefire* esteve associada às reduções significativas na vitimização de jovens em Boston.

O programa de prevenção e controle da criminalidade “Fica Vivo”, implementado na cidade de Belo Horizonte, Minas Gerais, foi baseado na intervenção *CeaseFire* que atua sob os princípios da teoria da desorganização social, desenvolvendo atividades que visam à mudança de comportamento nas comunidades de modo a reduzir, principalmente, os homicídios. O objetivo do Fica Vivo é reduzir os homicídios em áreas onde sua concentração é alta, em geral, favelas (*hot spots*). Para isto, combina ações preventivas e repressivas. Matta e Andrade (2005) avaliaram o Programa nos setores censitários pertencentes à região do Morro das Pedras. Por meio da metodologia de diferenças em diferenças, os resultados do estudo mostraram que, em média, a criminalidade nessa região reduziu 2,7% em relação às demais localidades de Belo Horizonte no período de 2000 a 2004.

Peixoto *et al.* (2007) também avaliaram o Fica Vivo considerando as sete áreas do município de Belo Horizonte onde o Programa foi implementado, até o ano de 2005 - Morro das Pedras, Pedreira Prado Lopes, Alto Vera Cruz, Taquaril, Cabana de Pai Tomás, Ribeiro de Abreu e Conjunto Felicidade. Para empreender tal avaliação, os autores aplicaram o método de diferenças em diferenças para dados semestrais entre os anos de 2000 e 2006, dividindo a análise em quatro períodos não uniformes, de acordo com as datas de expansão do Programa para as diferentes áreas, e para dois grupos de controle, todo o município de Belo Horizonte e somente as favelas deste município. Os resultados apontaram que o impacto do Fica Vivo é diferente nas áreas analisadas, sendo que o maior impacto do Programa foi observado na área piloto de implementação, Morro das Pedras.

Já Peixoto *et al.* (2008) promoveram uma avaliação econômica do Programa Fica Vivo no aglomerado do Morro das Pedras em Belo Horizonte, efetuando análises de custo-efetividade e de custo-benefício, entre os anos de 2004 e 2006. Os custos foram apurados pelo método de contagem, enquanto a efetividade foi mensurada pelo método diferenças em diferenças com pareamento por escore de propensão. Os resultados mostraram que o custo de um homicídio evitado pelo Programa (custo-efetividade) é de, aproximadamente, R\$ 244,6 mil, o que implica em uma taxa de retorno de, aproximadamente, 99%, ou seja, cada um real investido no Programa gera, no futuro, uma economia de R\$ 1,99 (custo-benefício).

Silveira *et al.* (2010) avaliaram o impacto do Fica Vivo por intermédio do modelo de regressão logística para comparar a incidência de homicídios no Morro das Pedras com aquela que ocorria ao mesmo tempo nas favelas não violentas, que não possuíam o Programa e com a incidência nos demais bairros da cidade de Belo Horizonte no período compreendido entre os anos de 2002 a 2006. Segundo conclusões do estudo, o Fica Vivo no Morro das Pedras foi responsável por reduções importantes e sustentadas dos homicídios desde sua implementação em 2002, de maneira que o Programa pode constituir em uma importante alternativa de prevenção de homicídios.

Muito se tem discutido sobre a relação entre armas e violência de maneira que vários tipos de políticas relativas à posse de armas de fogo têm sido adotados com vistas à redução da criminalidade. Podem ser citadas, as políticas de *buy back* (recompra de armas), proibição de posse de armas e permissão de porte de arma oculta. Nesse sentido, Lott Jr. e Mustard (1997) e Lott Jr. (2010) afirmam que a permissão do porte de armas pode reduzir a violência em decorrência da queda da probabilidade de sucesso do criminoso, visto que haveria um número maior de potenciais vítimas armadas. Entretanto, Duggan (2001), ao examinar a relação entre posse de armas e homicídios nos Estados Unidos nas décadas 1980 e 1990, chegou à conclusão de que mais armas disponíveis aumentaram a probabilidade de os conflitos interpessoais resultarem em morte.

Com o mesmo intuito de verificar a relação entre armas e violência, Villaveces *et al.* (2000) avaliaram o impacto da política DESEPAZ (Programa para Desenvolvimento, Segurança e Paz) que, dentre outras medidas, estabelecia a proibição de porte de armas de fogo em dias historicamente relacionados a altas taxas de homicídios, tais como feriados e dias de pagamento, nas cidades de Cali, cidade pioneira na adoção do DESEPAZ, e Bogotá, cidade que adotou política assemelhada, ambas cidades da Colômbia. O estudo utilizou séries de tempo interrompidas na qual Cali obteve 34 intervenções para um total de 89 dias, enquanto, em Bogotá, foram 22 intervenções para um total de 67 dias em três intervalos distintos (dezembro

de 1995 a março de 1996; dezembro de 1996 a fevereiro de 1997; março e abril de 1997). A partir da metodologia implementada com múltiplas replicações, os autores compararam períodos com a intervenção e períodos semelhantes sem a intervenção. O principal resultado do estudo aponta que, em ambas as cidades, o Programa reduziu a ocorrência de homicídios por armas de fogo. Entretanto, o mesmo impacto é observado em outros tipos de homicídios, de forma que não ficou suficientemente claro o mecanismo pelo qual a intervenção atua sobre a violência: se pela proibição do porte de armas ou pela intensificação do policiamento.

No Brasil, Souza *et al.* (2007), por intermédio de uma série histórica de óbitos por armas de fogo, calcularam o número estimado de óbitos para o primeiro semestre e segundo semestre de 2004, períodos em que foram implementados o Estatuto do Desarmamento (ED) e a campanha nacional de entrega de armas, respectivamente. Com isso, os autores encontraram evidências de que a redução das mortes e hospitalizações por armas de fogo pode ser atribuída às medidas tomadas pelo governo federal para reduzir a disponibilidade de armas à população. Deste modo, os autores estimaram 5.563 vidas poupadas, sendo 2.292 no primeiro e 3.271 no segundo semestre.

Hartung (2009) analisou o impacto das políticas públicas do governo estadual de repressão ao porte ilegal de armas e o ED sobre as taxas de homicídios nos 645 municípios de São Paulo no período de 1999 a 2007. Por intermédio de duas diferentes abordagens metodológicas, vetor autorregressivo e painel dinâmico, o autor conclui que a redução no estoque de armas foi um fator crucial para a redução de homicídios no estado, sugerindo que o ED pode ter sido uma importante política pública para reduzir a criminalidade letal.

Cerqueira (2010) e Cerqueira e Mello (2012) investigaram o impacto do controle das armas de fogo sobre diversos tipos de crime (homicídios dolosos, latrocínios, lesão corporal dolosa, roubos de veículos, crimes associados a drogas ilícitas) para os 645 municípios do estado de São Paulo entre 2001 e 2007. Para tanto, o autor utilizou o ED como a fonte de variação exógena na disponibilidade de armas, atentando-se para o fato de que efeito do ED poderia ser maior nos municípios com maior prevalência de armas e/ou com maior densidade populacional. Com isso, a variável endógena estoque de armas foi instrumentalizada pela *dummy* de ED interagida com a prevalência de armas nos municípios em 2003, captada pela proporção de suicídios por armas de fogo no mesmo ano, e pela interação entre o ED e o tamanho das cidades (variável categórica de classificação dos municípios segundo a população residente). Os resultados a partir das estimações por mínimos quadrados em dois estágios (MQ2E) com erro-padrão *cluster* robusto, evidenciaram que, no período analisado, houve efetivamente uma diminuição na prevalência de armas de fogo no Estado de São Paulo e que o

ED gerou efeitos importantes para a redução da taxa de homicídios dolosos e por armas de fogo. Contudo, o ED não impactou significativamente os crimes contra o patrimônio (latrocínio e de roubo de veículos).

Em Justus e Kassouf (2012), a avaliação do ED foi realizada por meio de uma análise de intervenção baseada em Box e Tiao (1975) com dados trimestrais das taxas de crimes letais para a cidade de São Paulo. Os autores concluíram que o ED reduziu as taxas de criminalidade violenta na capital do estado. Já no estudo de Justus e Kassouf (2013), os resultados do modelo vetorial autorregressivo com correção de erro trazem evidências estatísticas de que a redução da taxa de desemprego, a maior eficiência da polícia e o ED foram fatores preponderantes para explicar a queda das taxas de crimes da cidade de São Paulo.

Já Scorzafave *et al.* (2015) avaliaram o impacto da campanha de entrega voluntária de armas de fogo sobre a incidência de óbitos com armas de fogo no estado do Paraná, usando o estimador de diferenças em diferenças. Pelo fato de a campanha pelo desarmamento ter iniciado seis meses antes da campanha nacional no Paraná, os municípios paranaenses constituíram o grupo de tratamento e os municípios das regiões Centro-Oeste, Sudeste e Sul do País, delimitados por meio de técnicas de pareamento (*propensity score matching*), formaram o grupo de controle. Os autores concluíram que a política de *buy back* não contribuiu, pelo menos no curto prazo, para a redução dos crimes envolvendo armas de fogo no estado do Paraná.

Em relação aos trabalhos empíricos, percebe-se que diversas abordagens metodológicas são adotadas para a avaliação dos mais diferentes tipos de políticas de segurança pública. Essa constante avaliação de programas demonstra a importância do tema, haja visto que a garantia da segurança pública tem impactos diretos sobre o bem-estar da sociedade e sobre o desenvolvimento econômico. O Quadro 2.1 resume alguns trabalhos sobre avaliação dos diferentes tipos de políticas de segurança pública, mantendo uma ordem dos estudos em níveis internacional e nacional, com enfoque sobre os estudos realizados com vistas à compreensão das causas para a trajetória declinante da criminalidade no estado de São Paulo nos anos 2000.

Quadro 2.1 - Resumo dos trabalhos empíricos sobre avaliação de políticas de segurança pública.

(continua)

Autores	Região	Período	Política analisada	Método utilizado	Controle Espacial	Variável dependente	Variáveis explicativas	Conclusões
Braga <i>et al.</i> (2001)	Boston / EUA	Mensal: jan/1991 a mai/1998	Policimento orientado para os problemas: Operação <i>Ceasefire</i>	Regressão Poisson em um modelo log-linear	Não	Homicídios de jovens de 14 a 24 anos	<i>Dummy</i> da Operação <i>Ceasefire</i> (maio de 1996), termos de tendências e <i>dummies</i> temporais	A Operação <i>Ceasefire</i> esteve associada a reduções significativas de homicídios de jovens em Boston.
Kelling e Souza (2001)	Nova Iorque / EUA	Anual: 1989-1998 e 2000	<i>Broken Windows</i> e COMPSTAT	Modelo hierárquico linear e estudo de caso	Não	Homicídios, estupros, latrocínios e roubos	Desemprego, população, prisões por contravenção e uso de drogas	O policiamento “tolerância zero” reduziu os crimes violentos e o uso do COMPSTAT é um importante para a definição de estratégias de combate ao crime.
Levitt (2004)	EUA	Anual: 1973-1991 e 1991-2001	<i>Broken Windows</i> , COMPSTAT e policiamento comunitário	Estudo qualitativo	Não	Homicídios, estupros, furto, latrocínios, roubos	PIB <i>per capita</i> , envelhecimento da população, diferentes políticas, controle de armas, escolaridade, efetivo policial, prisões, uso de droga (crack), legalização do aborto	O aumento do efetivo policial e das prisões, a redução do uso de droga (crack) e a legalização do aborto em 1973 parecem ser os fatores que explicaram o declínio da criminalidade nos EUA verificada nos anos 1990.
Rosenfeld <i>et al.</i> (2005)	Boston, Nova Iorque e Richmond na Virgínia / EUA	Anual: 1992-2001	1. COMPSTAT, 2. <i>Ceasefire</i> e 3. <i>Exile</i>	Curvas de crescimento e modelos hierárquicos lineares generalizados	Não	1. Homicídios de jovens por arma de fogo 2. Homicídios por armas de fogo	Densidade demográfica, efetivo policial, taxa de prisão, prevalência de cocaína, <i>dummy</i> para cada intervenção	A tendência de homicídios da cidade de Nova Iorque na década de 1990 não diferiram de forma significativa em relação a outras grandes cidades. A queda mais acentuada nos homicídios ocorreu em Boston.

Quadro 2.1 - Resumo dos trabalhos empíricos sobre avaliação de políticas de segurança pública.

(continuação)

Autores	Região	Período	Política analisada	Método utilizado	Controle Espacial	Variável dependente	Variáveis explicativas	Conclusões
Corman e Mocan (2005)	Cidade de Nova Iorque	Mensal: dez/1974 a dez/1999	<i>Broken Windows</i>	Séries temporais	Não	Homicídio, assalto, roubo, furto, roubo de veículos, estupro	Taxa de desemprego, salário mínimo, prisões, efetivo policial, número de encarcerados, jovens, efeito sazonal do crime	Medidas de dissuasão tem impacto maior do que variáveis econômicas para a queda da criminalidade. As táticas repressivas de policiamento (prisões por contravenção – <i>proxy</i> do “tolerância zero”) reduziram furtos e roubos de veículos
Mazerolle <i>et al.</i> (2007)	Queensland, Austrália	Mensal: jul/1995 a jun/2004	<i>Operational Performance Reviews (OPRs)</i> - Similar ao COMPSTAT	Séries temporais	Não	13 diferentes tipos de crimes	Processos autorregressivos e de médias móveis	A introdução das OPRs reduziu os crimes, principalmente, roubos a residências e outros crimes contra o patrimônio. Ademais, as OPRs parecem ser custo-efetivas no controle ao crime, pois resultaram em uma economia de AUD\$ 1.162.175,00.
Weisburd <i>et al.</i> (2008)	Diferentes cidades dos EUA	—	Policiamento orientado para os problemas (POP)	Revisão de literatura sobre diversos estudos que analisaram POP	Não	Homicídio, assalto, roubo, furto, roubo de veículos, estupro	Efetivo policial, uso de drogas lícitas e ilícitas, prisões	O policiamento orientado para os problemas reduz o crime, apesar de seu efeito ser modesto.
Jang <i>et al.</i> (2010)	Texas / EUA	Mensal: jan/2000 a nov/2006	<i>All Staff Meeting (ASM)</i> - Similar ao COMPSTAT	Séries temporais: análise de intervenção e multivariada	Não	Embriaguez, Drogas, Vandalismo, Prostituição; Taxas de homicídio, roubo, furto, roubo de veículos	<i>Dummy</i> indicadora do COMPSTAT (set/2002), desemprego, homens jovens, heterogeneidade racial, mobilidade residencial, lares uniparentais,	As prisões “tolerância zero” aumentaram após o COMPSTAT. Houve queda nos índices de crime total e nos crimes contra a propriedade. Entretanto, não houve mudança significativa nos crimes violentos.

Quadro 2.1 - Resumo dos trabalhos empíricos sobre avaliação de políticas de segurança pública.

(continuação)

Autores	Região	Período	Política analisada	Método utilizado	Controle Espacial	Variável dependente	Variáveis explicativas	Conclusões
Matta e Andrade (2005)	Setores censitários de Belo Horizonte (MG)	Anual: 2000-2004	Policiamento orientado para os problemas: Fica Vivo	DID	Não	Número de homicídios	População jovem masculina, baixa renda, anos de escolaridade	A criminalidade nos setores censitários contidos na região do Morro das Pedras reduziu em 2,7% em relação aos demais setores censitários.
Peixoto <i>et al.</i> (2007)	Sete favelas do município de BH (MG)	Semestral: S1/2000 a S2/2006	Policiamento orientado para os problemas: Fica Vivo	DID	Não	Taxa de homicídio/ 100 mil	Características do domicílio (água, esgoto, lixo) e do residente (alfabetização, quartil de renda, faixa etária)	O impacto do Fica Vivo é maior na área piloto de implementação, Morro das Pedras.
Ferreira (2013)	DPs dos municípios do Rio de Janeiro	Anual: 2005 a 2012	Policiamento comunitário: UPPs	DID	Não	Taxa de homicídio/ 100 mil, latrocínios e roubos a veículos	População e área geográfica de cada DP	Os homicídios diminuíram tanto na Baixada Fluminense quanto na Capital, porém o Programa não alterou os índices de latrocínios cometidos no estado nem tampouco reduziu os índices de roubo a veículos na região da Baixada.
Suliano e Oliveira (2013)	Bairros da RM de Fortaleza (CE)	Mensal: jan-mai de 2007 e jan-maio de 2008	Policiamento comunitário: Ronda do Quarteirão	DID	Não	Taxa de roubos e de furtos por mil	Renda, desemprego, população com ensino médio completo, lares uniparentais, densidade, <i>dummy</i> de intervenção	O aumento recente do efetivo policial nos bairros tratados apresentou uma redução de 57% nas taxas de roubos.
Silveira Neto <i>et al.</i> (2013)	Estado de Pernambuco (PE)	Anual: 2007 a 2011	Pacto pela Vida – <i>hot spots</i>	Comparação PE e o grupo de acordo com o controle sintético	Não	Taxa de homicídio/ 100 mil	Renda, desigualdade, desemprego, densidade, escala populacional e população jovem	Houve uma redução anual de 9,1 pontos na taxa de homicídio de Pernambuco quando comparado ao seu controle sintético para o período 2007-2011.

Quadro 2.1 - Resumo dos trabalhos empíricos sobre avaliação de políticas de segurança pública

(continuação)

Autores	Região	Período	Política analisada	Método utilizado	Controle Espacial	Variável dependente	Variáveis explicativas	Conclusões
Kahn e Zanetic (2005)	39 municípios da RMSP	Anual: 2000-2004; 1996-2004 Mensal: jan/1998 a mar/2005	ED; Guarda Municipal, Disque Denúncia; Secretaria de Segurança Pública Municipal	Estatísticas descritivas e regressão simples	Não	Homicídios; agressões por armas de fogo; crimes contra o patrimônio	Número de suicídios por armas de fogo, atendimentos por intoxicação por álcool, religião	A divulgação do Disque Denúncia e a Lei Seca contribuem para diminuir a quantidade de homicídios. Já a guarda municipal pode contribuir para a redução dos crimes contra o patrimônio.
Scripilliti (2006)	645 municípios paulistas	Anual: 1997 e 2003	Disque Denúncia, Restrição ao consumo de álcool, Guardas Municipais e SSP municipal	DID	Não	Crimes contra patrimônio e homicídios dolosos	Densidade, urbanização, despesas (total, educação, habitação, transporte, com pessoal, previdência), população, emprego (total e industrial), IPRS (renda, longevidade, educação)	As diversas políticas analisadas surtiram efeito na redução dos homicídios dolosos. Entretanto, não demonstraram efeito significativo para a redução dos crimes contra o patrimônio.
Goertzel e Kahn (2007)	Cidade e estado de São Paulo	Anual: 2001 a 2007	Policimento padrão e <i>hot spot</i>	Análise descritiva	Não	Homicídios	Atividade policial (prisões), restrição de armas	A redução da criminalidade foi creditada ao policiamento mais eficaz, incluindo o controle de armas.
Hartung (2009)	645 municípios paulistas	1999 a 2007	<i>Buy back</i> : (ED)	VAR e painel dinâmico	Não	Taxa de homicídios/100 mil	Homicídios _{t-1} , jovens, urbanização, população total, desemprego, salário, Infocrim, crimes solucionados, policiais e prisões <i>per capita</i>	Em ambas as estratégias econométricas, a redução no estoque de armas foi um fator crucial para a redução de homicídios no estado

Quadro 2.1 - Resumo dos trabalhos empíricos sobre avaliação de políticas de segurança pública

(conclusão)

Autores	Região	Período	Política analisada	Método utilizado	Controle Espacial	Variável dependente	Variáveis explicativas	Conclusões
Dias (2011)	Estado de São Paulo	Anual: 1990 a 2010	“Hipótese PCC”	Estatísticas descritivas	Não	Homicídios	População carcerária, rebeliões, fugas e apreensões de presos, estrutura PCC.	A monopolização do “mundo do crime” no sistema carcerário paulista explica a drástica redução do uso da violência física nos conflitos interpessoais.
Cerqueira e Mello (2012)	645 municípios paulistas	Anual: 2001 a 2007	<i>Buy back</i> : (ED)	MQ2E com erro-padrão <i>cluster</i> robusto	Não	Taxas/100 mil: homicídios, latrocínios, lesão corporal, roubo de veículos, crimes associados a drogas ilícitas	Interação entre ED e prevalência de armas nos municípios em 2003, interação entre o ED e o tamanho das cidades, <i>dummies</i> temporais	O ED foi um dos fatores que contribuíram para a redução nos crimes violentos. Não há evidências de qualquer efeito sobre outros crimes econômicos.
Justus e Kassouf (2012)	Cidade de São Paulo	Trimestral: 1995T3 a 2010T4	<i>Buy back</i> : (ED)	Séries de tempo: análise de intervenção	Não	Taxa de crimes letais (homicídios e latrocínios) /100 mil	<i>Dummy</i> de intervenção (ED); defasagens temporais (AR3; MA1)	O ED causou um efeito negativo na taxa de crimes letais.
Justus e Kassouf (2013)	Cidade de São Paulo	Trimestral: 1997T1 a 2010T4	<i>Buy back</i> : (ED)	Séries de tempo: VEC cointegração	Não	Taxa de crimes letais (homicídios e latrocínios) /100 mil	Salário, desemprego, armas apreendidas/ revistas, crime por tráfico/prisões, ED.	Maiores salários, melhor desempenho da polícia e o ED são fatores de redução do crime enquanto o desemprego tem efeito positivo.
Willis (2013)	Cidade de São Paulo	Anual: 2007 a 2012	“Hipótese PCC”	Estudo de caso	Não	Homicídios	Rotina de policiais do DHPP, documentos internos apreendidos com um membro do PCC, entrevista com moradores.	A consolidação e regulação do PCC é o principal fator que explica a redução dos homicídios na cidade de São Paulo. O PCC é produto, produtor e regulador da violência.

Fonte: Elaboração Própria.

Com base no Quadro 2.1, percebe-se que a avaliação de políticas de segurança pública no Brasil é um tema recente e ainda pouco explorado. Especificamente no que se refere à análise da queda vertiginosa dos índices de criminalidade em São Paulo na primeira década dos anos 2000, nota-se que os estudos que promoveram uma avaliação de política de segurança pública se concentraram basicamente no impacto da política federal de restrição ao uso de armas de fogo, o Estatuto do Desarmamento. Sendo assim, essa tese contribui para a literatura de economia do crime, mais especificamente, para a avaliação de política ao analisar o impacto da política de segurança pública específica de São Paulo – INFOCRIM – sobre a criminalidade, controlando por diversos determinantes do crime e para outras políticas de segurança pública implementadas no período.

Apesar do recente interesse dos pesquisadores em avaliar o impacto das diferentes políticas de segurança pública, Santos (2013) observa que o deslocamento de *hot spots* pós-intervenção pode ter efeito redutor sobre o impacto da política de combate à criminalidade empreendida. Nessa mesma linha, Matta e Andrade (2005) alertam para o fato de que se deve ter o devido cuidado ao afirmar sobre o impacto de um Programa específico no combate à criminalidade, uma vez que um dos motivos para a redução da criminalidade seria a possibilidade de migração do crime da região de tratamento para outras regiões vizinhas, em busca de um ambiente mais favorável à atividade criminal. Em Justus e Santos Filho (2011), por exemplo, foram encontradas evidências de efeitos de transbordamento (*spillovers effects*) entre as taxas de crime das microrregiões brasileiras.

Considerando essas ressalvas, as interações espaciais podem ser relevantes para explicar o efeito médio do tratamento, uma vez que, caso existam, potencializam ou suavizam o impacto da política implementada. Diante disso, outra contribuição desse trabalho consiste em propor o uso da metodologia de diferenças em diferenças espacial (SDID) capaz de modelar a possível migração do crime das regiões tratadas para regiões vizinhas, desconsiderada até então nas avaliações de impacto de políticas de segurança pública.

3 METODOLOGIA

A literatura microeconômica de avaliação de tratamento visa mensurar o impacto de mudanças aleatórias, denominadas experimentos naturais, sobre o grupo afetado pela intervenção⁶. Em outras palavras, a avaliação de programas busca captar o efeito de algum evento exógeno, como, por exemplo, uma alteração institucional, mudança nas leis ou alteração nas políticas governamentais sobre o grupo experimental (WOOLDRIDGE, 2010).

A característica central da metodologia experimental envolve um experimento puro em que há a aleatorização da amostra e os mesmos indivíduos são observados nos dois estados da natureza (tratado e não tratado). A distribuição aleatória implica que os indivíduos expostos ao tratamento são escolhidos aleatoriamente e, portanto, a designação do tratamento não depende do resultado e é não correlacionado com os atributos dos indivíduos tratados. Com isso, uma forma de mensurar o impacto do tratamento seria estimar a diferença da média da variável de resultado do mesmo indivíduo nos dois estados da natureza, tratado e não tratado (CAMERON e TRIVEDI, 2005).

Entretanto, tais condições não podem ser encontradas na maioria das intervenções nas ciências sociais, levando assim a uma possível falha de identificação dos parâmetros causais de interesse pelo fato de a configuração aleatória não ser factível. Normalmente, estas intervenções têm desenho não aleatório e as avaliações utilizam desenhos não experimentais (ANGRIST e KRUEGER, 1998).

Contudo, conforme aponta Cameron e Trivedi (2005), às vezes, as condições experimentais podem ser aproximadamente replicadas nos dados observacionais, criando as condições de um estudo quase-experimental, no qual o grupo tratado pode ser comparado com o grupo de controle. Esta avaliação de tratamento baseada em dados observacionais é conhecida como avaliações de estudos observacionais ou quase-experimentais (MEYER, 1995; CAMERON e TRIVEDI, 2005).

Dessa forma, pode-se estimar o efeito da intervenção por meio da comparação entre os resultados do grupo experimental submetido a um tratamento (grupo de tratamento), com indivíduos que possuem características idênticas e, portanto, elegíveis

⁶ “Em aplicações econômicas, tratamento e intervenção geralmente significam a mesma coisa” (CAMERON e TRIVEDI, 2005, p.260). Nesta tese, tais termos serão considerados sinônimos.

ao tratamento, porém não tratados (grupo de controle). Sendo assim, para empreender uma avaliação quase-experimental apropriada, devido cuidado deve ser exercido para compatibilizar de forma adequada os grupos de tratamento e de controle de modo a evitar possíveis vieses nas estimativas do impacto do tratamento.

Tendo em vista que a intervenção pode impactar o bem-estar da sociedade, a avaliação de políticas e programas tornou-se relevante para as funções de planejamento e gestão, bem como para a divulgação de resultados da política. Assim, a adoção de métodos de avaliação de intervenção em áreas prioritárias, como a segurança pública, torna-se importante para apoiar, subsidiar e apontar para os *policy makers* quais os tipos de políticas, as regiões e as áreas que necessitam ser atendidas por programas governamentais.

Diante disso, a avaliação de políticas tem atraído a atenção de diversos pesquisadores que buscam estimar o resultado das políticas. O método amplamente empregado para estimar tal resultado é o método de diferenças em diferenças (DID), a ser descrito na seção 3.1. Esse método requer diversos pressupostos para estimar de forma consistente e eficiente o impacto das políticas. Entretanto, tais premissas podem ser difíceis de serem sustentadas em diversas áreas, como no caso de uma implementação de política de segurança pública devido à possibilidade de migração do crime. Para contornar esse impasse, essa tese propõe avaliar o programa de segurança pública – INFOCRIM – implementado pelo Governo do Estado de São Paulo, por intermédio do método diferenças em diferenças espacial (SDID), descrito de forma pormenorizada na seção 3.2, de maneira a identificar corretamente o efeito da política.

Este capítulo apresenta também o modelo sob a ótica da teoria econômica do crime na seção 3.3 com vistas a dar argumento teórico à estratégia empírica adotada para estimar o impacto do INFOCRIM descrita na seção 3.4. Na referida seção, são apresentadas, ainda, a descrição e as estatísticas descritivas das variáveis utilizadas. Com o intuito de dar informações quantitativas à população acerca da criminalidade violenta evitada, a seção 3.5 explicita a metodologia utilizada para mensurar o número de vidas poupadas pelo INFOCRIM. E, finalmente, por se tratarem de dados agrupados em nível municipal, na seção 3.6 será realizada uma análise exploratória dos dados espaciais (AEDE).

3.1 DIFERENÇAS EM DIFERENÇAS⁷

Em um experimento natural ou quase-experimental, um dos métodos para a estimação dos efeitos do tratamento é um método de regressão simples baseado na comparação da média dos resultados dos grupos de controle e tratamento antes e depois da intervenção, conhecido na literatura como diferenças em diferenças.

De modo a formalizar o método DID, seja N o número de indivíduos elegíveis ao tratamento, N_T o número de indivíduos que receberam o tratamento e $N_C = N - N_T$ o número de indivíduos não tratados que servem como potencial grupo de controle. Considere, ainda, $(y_i, \mathbf{X}_i, D_i; i = 1, \dots, N)$ em que y é a variável de resultado, \mathbf{X} são as variáveis de controle e D refere-se a uma variável *dummy* que assume valor um (1) para os indivíduos tratados e zero (0), caso contrário. A variável de resultado do indivíduo tratado é denotada como y_1 ($y_1 = y_1^1 - y_1^0$) e para o indivíduo não tratado é denotada por y_0 ($y_0 = y_0^1 - y_0^0$), em que os sobrescritos 1 e 0 referem-se ao período de tempo depois e antes da intervenção, respectivamente.

Comparações significativas entre os resultados dos grupos de tratamento e de controle requerem algumas pressuposições de modo a permitir a identificação dos efeitos causais da política de intervenção por intermédio do método DID. As premissas fundamentais são descritas a seguir:

- i) Pressuposição de equilíbrio parcial: a premissa de equilíbrio parcial afirma que a avaliação de tratamento assume ausência de efeito de equilíbrio geral, ou seja, os efeitos do tratamento são pequenos e não afetam as variáveis consideradas como exógenas.
- ii) Pressuposição de *stable unit treatment value assumption* (SUTVA): estabelecida por Rubin (1974), essa premissa assume a hipótese de que o tratamento não afete indiretamente as observações não tratadas de modo a identificar o parâmetro do impacto da intervenção.
- iii) Pressuposição de independência condicional: Estabelece que, condicional a \mathbf{X} , os resultados dos grupos de tratamento e de controle são independentes do tratamento, isto é:

$$y_0, y_1 \perp D | \mathbf{X} \tag{3.1}$$

⁷ Esta seção baseia-se em Meyer (1995), Cameron e Trivedi (2005) e Wooldridge (2010).

A implicação dessa pressuposição é a de que a participação no tratamento não depende dos resultados, após controlar para as características observáveis \mathbf{X} . Se válida, essa premissa estabelece que, uma vez controladas as características observáveis \mathbf{X} , algumas das quais podem estar relacionadas a D , o tratamento e o resultado são independentes e será útil para a identificação do parâmetro do impacto da intervenção.

Uma pressuposição mais fraca que (3.1) é a de independência condicional de participação, ou seja, y_0 não é afetado pelo tratamento. Em termos formais:

$$y_0 \perp D | \mathbf{X} \quad (3.2)$$

Essa pressuposição é conhecida como pressuposição de ignorabilidade.⁸ Se (3.2) é válida, implica que não existe viés de variável relevante observável omitida, uma vez que \mathbf{X} é incluída na regressão.

iv) Pressuposição de média condicional: a premissa de independência da média condicional implica que a média de y_0 não determina a participação no programa.

$$E[y_0 | D = 1, \mathbf{X}] = E[y_0 | D = 0, \mathbf{X}] = E[y_0 | \mathbf{X}] \quad (3.3)$$

De outra forma, a média do resultado dos indivíduos não tratados está condicionada somente às suas características observáveis.

Garantidas as premissas fundamentais subjacentes à avaliação do tratamento, o método DID pode ser empregado para estimar o efeito médio do tratamento. Seguindo a notação de Meyer (1995), o método DID pode ser expresso no contexto do modelo de regressão linear:

$$y_{ij}^t = \alpha + \beta_1 D^t + \beta_2 D_j + \phi D_j^t + \varepsilon_{ij}^t \quad (3.4)$$

em que i denota os indivíduos elegíveis ao tratamento, j é o subscrito para cada grupo em que $j = 0$ refere-se ao grupo de controle e $j = 1$ é o grupo de tratamento, t é o tempo onde $t = 0$ indica o período pré-intervenção e $t = 1$ ao período pós-intervenção. A

⁸ Esse pressuposto é também conhecido como seleção nos observáveis ou não-confundimento (*unconfoundedness assumption*).

variável *dummy* D^t representa uma variável binária temporal que se refere aos efeitos específicos do tempo, sendo $D^t = 1$, o período pós tratamento e 0, caso contrário. A variável *dummy* D_j é definida como sendo $D_j = 1$ se o indivíduo faz parte do grupo de tratamento e zero (0), caso o indivíduo pertença ao grupo de controle. Os parâmetros α e β_1 são coeficientes a serem estimados em que α capta o efeito fixo no tempo e β_1 é o efeito específico do tempo, capturando os efeitos comuns e homogêneos a todos os indivíduos, tratados ou não.

Já a variável D_j^t é uma variável *dummy* resultante da interação entre D^t e D_j que assume valor 1 para indicar o indivíduo tratado no período pós-intervenção e 0, caso contrário. O coeficiente \emptyset da *dummy* D_j^t a ser estimado é o parâmetro de interesse que fornece o efeito do tratamento. O termo de erro ε_{ij}^t é, por suposição, um termo de erro aleatório, ou seja, possui média zero e variância constante. A equação (3.4) pré (*ex ante*) e pós (*ex post*) intervenção para o grupo de tratamento passa a ser, respectivamente:

$$y_{i1}^0 = \alpha + \beta_2 + \varepsilon_{i1}^0 \quad (3.5)$$

$$y_{i1}^1 = \alpha + \beta_1 + \beta_2 + \emptyset + \varepsilon_{i1}^1 \quad (3.6)$$

O efeito médio do tratamento sobre o grupo tratado (ATET⁹) pode ser obtido por meio da aplicação do operador esperança sobre a diferença entre os períodos pós e pré-intervenção, como segue:

$$E[y_{i1}^1 - y_{i1}^0] = \beta_1 + \emptyset + E[\varepsilon_{i1}^1 - \varepsilon_{i1}^0] \quad (3.7)$$

Dada a suposição de que ε_{ij}^t é um termo de erro aleatório, $E[(\varepsilon_{i1}^1 - \varepsilon_{i1}^0)] = 0$, o efeito médio da intervenção sobre o grupo tratado passa a ser:

$$E[y_{i1}^1 - y_{i1}^0] = \beta_1 + \emptyset \quad (3.8)$$

O ATET pode ser estimado de forma simples por meio do método de mínimos quadrados ordinários (MQO) por intermédio da equação descrita abaixo:

⁹ ATET é acrônimo de *average treatment effect on the treated*. Em livre tradução, efeito médio do tratamento sobre o tratado.

$$y_1 = \beta_1 + \emptyset D_1^1 + v_1 \quad (3.9)$$

Onde $v_1 = \Delta\varepsilon_1 = \varepsilon_{i1}^1 - \varepsilon_{i1}^0$ representa um termo de erro aleatório. Incluindo variáveis explicativas (\mathbf{X}) na equação (3.9) de modo a controlá-la para as características observáveis com seus respectivos parâmetros a serem estimados ($\boldsymbol{\theta}$), tem-se:

$$y_1 = \beta_1 + \emptyset D_1^1 + \Delta\mathbf{X}\boldsymbol{\theta} + v_1 \quad (3.10)$$

De forma análoga, o efeito temporal médio no intervalo pós e pré-intervenção sobre o grupo de controle pode ser expresso da seguinte forma:

$$E[y_{i0}^1 - y_{i0}^0] = \beta_1 + E[\varepsilon_{i0}^1 - \varepsilon_{i0}^0] \quad (3.11)$$

Rearranjando a equação (3.11), tem-se:

$$y_0 = \beta_1 + v_0 \quad (3.12)$$

Em que $v_0 = \Delta\varepsilon_0 = \varepsilon_{i0}^1 - \varepsilon_{i0}^0$ é um termo de erro aleatório. Controlando para as características observáveis, a equação (3.12) pode ser reescrita como segue:

$$y_0 = \beta_1 + \Delta\mathbf{X}\boldsymbol{\theta} + v_0 \quad (3.13)$$

O efeito da intervenção pode ser calculado a partir da subtração da dupla diferença, conforme mostrado a seguir:

$$(y_{i1}^1 - y_{i1}^0) - (y_{i0}^1 - y_{i0}^0) = \emptyset + (\varepsilon_{i1}^1 - \varepsilon_{i1}^0) - (\varepsilon_{i0}^1 - \varepsilon_{i0}^0) \quad (3.14)$$

Aplicando o operador de esperança em (3.14), a equação toma a seguinte forma:

$$E[(y_{i1}^1 - y_{i1}^0) - (y_{i0}^1 - y_{i0}^0)] = E[\emptyset] + E[(\varepsilon_{i1}^1 - \varepsilon_{i1}^0) - (\varepsilon_{i0}^1 - \varepsilon_{i0}^0)] \quad (3.15)$$

Dada a pressuposição que ε_{ij}^t é um termo de erro aleatório, $E[(\varepsilon_{i1}^1 - \varepsilon_{i1}^0) - (\varepsilon_{i0}^1 - \varepsilon_{i0}^0)] = 0$, é possível obter o efeito médio do tratamento como segue:

$$E[(y_{i1}^1 - y_{i1}^0) - (y_{i0}^1 - y_{i0}^0)] = \emptyset \quad (3.16)$$

A equação (3.16) é conhecida na literatura como o estimador DID e calcula o efeito médio do tratamento (ATE¹⁰). O ATE, garantidas todas as premissas, pode ser estimado pelo método de MQO através da especificação geral do DID a seguir:

$$\Delta y_i = \beta_1 + \emptyset D_j^1 + \Delta X_i \theta + v_i \quad (3.17)$$

O estimador DID tem a vantagem de eliminar o problema relativo à omissão de variável relevante invariante no tempo e alguns possíveis problemas relacionados à forma funcional (McMILLEN, 2010). Além disso, o estimador DID também é capaz de controlar os efeitos não observados idiossincráticos invariantes no tempo.

Em resumo, o estimador de diferenças em diferenças possibilita isolar o efeito do grupo de tratamento de maneira simples e objetiva, ou seja, o DID possibilita separar o efeito da política de interesse sobre a variável dependente, controlando para todos outros possíveis fatores que explicam a variável dependente. O Quadro 3.1 resume o estimador do método DID.

Quadro 3.1 – Resumo esquemático do método DID.

Grupo/Período	<i>Ex ante</i>	<i>Ex post</i>	Diferenças
Tratamento	$\alpha + \beta_2$	$\alpha + \beta_1 + \beta_2 + \emptyset$	$\beta_1 + \emptyset$
Controle	α	$\alpha + \beta_1$	β_1
Diferenças	β_2	$\beta_2 + \emptyset$	\emptyset

Fonte: Elaboração Própria.

Como visto, o estimador DID é capaz de identificar o efeito causal do tratamento de forma eficiente e consistente, quando as premissas subjacentes ao DID são garantidas. Entretanto, quando a intervenção é implementada sobre indivíduos que são unidades geográficas (municípios, estados, regiões entre outros), a hipótese SUTVA pode não ser mais válida, uma vez que, pelo fato da possibilidade de interação espacial entre os

¹⁰ ATE é acrônimo de *average treatment effect*. Em livre tradução, efeito médio do tratamento.

agentes, pode ocasionar transbordamento espacial da intervenção do grupo de tratamento para o grupo de controle. Diante disso, o efeito do tratamento pode ser reforçado, ou atenuado, em virtude de tais interações espaciais entre os grupos de tratamento e de controle.

Caso haja evidências de autocorrelação espacial, a premissa de independência condicional também não pode ser sustentada e, diante disso, não será possível identificar o efeito causal da política. Sendo assim, deve-se adotar um método que controle adequadamente as interações espaciais entre os agentes para, então, mensurar o efeito médio do tratamento. O método diferenças em diferenças espacial (SDID) é o mais apropriado para esse caso.

3.2 DIFERENÇAS EM DIFERENÇAS ESPACIAL¹¹

Os cientistas regionais têm demonstrado a importância de incluir o espaço na modelagem de maneira a garantir estimativas consistentes e/ou eficientes, uma vez que o fato de ignorar as interações espaciais entre os agentes pode acarretar viés de variável relevante omitida prejudicando, por conseguinte, as inferências das análises realizadas. Da mesma forma, as interações espaciais podem ser relevantes para explicar o efeito médio do tratamento, uma vez que elas podem sobreidentificar ou subidentificar o impacto da intervenção.

Conforme argumentado por Dubé *et al.* (2014), DID é um método que controla adequadamente a estrutura e a interação espaciais constantes no tempo. Entretanto, a interação espacial entre os agentes, que varia ao longo do tempo, não é captada e necessita ser modelada no método DID para estimar o efeito causal do tratamento de forma consistente e eficiente.

Uma primeira tentativa de modelar o espaço em uma avaliação de tratamento é creditada a McIntosh (2008). O autor mostra que não controlar a heterogeneidade espacial viesava o efeito causal das políticas implementadas em unidades regionais. Para controlar a heterogeneidade espacial, McIntosh (2008) utilizou uma terceira diferenciação, subtraindo o efeito espacial não observado das estimativas do DID de modo a estimar o

¹¹ Essa subseção é baseada em Almeida (2014).

resultado do efeito do tratamento sobre o tratado. Há duas premissas subjacentes ao modelo proposto por McIntosh (2008): a primeira é que os indivíduos pertencentes aos grupos de tratamento e de controle localizados, no mesmo espaço físico, estão sujeitos à mesma influência da heterogeneidade espacial. Já a segunda premissa estabelece que não há efeitos *spillovers* da intervenção dos indivíduos do grupo de tratamento para os pertencentes ao grupo de controle dentro da mesma região mantendo, portanto, a hipótese SUTVA.

Com o mesmo objetivo de modelar heterogeneidade espacial, Heckert e Mennis (2012) estimaram dois modelos DID para diferentes unidades regionais e o modelo das regressões ponderadas geograficamente (RPG), evidenciando que a modelagem do espaço é importante para a avaliação de políticas.

Cabe salientar que, apesar de os estudos supracitados perceberem a importância do espaço nas avaliações de tratamento, os mesmos esforçaram-se em modelar apenas um dos efeitos espaciais, qual seja, a heterogeneidade espacial. Contudo, há de se atentar para o fato de que as políticas implementadas em unidades geográficas podem gerar a autocorrelação espacial que deriva teoricamente da interação entre os agentes agregados em unidades geográficas. Uma vez omitida nas avaliações de tratamento, a autocorrelação espacial pode viesar e, até mesmo, alterar o efeito causal do tratamento, culminando em inferências errôneas. Quanto à heterogeneidade espacial, derivada teoricamente da diversidade da estrutura (econômica, social, política, entre outras), esta torna as estimativas menos eficientes, porém consistentes, constituindo, portanto, um problema menos grave que a omissão da dependência espacial. Conforme relatado por Anselin (1988), esse tipo de problema costuma ser resolvido por técnicas tradicionais de econometria, com a inclusão de *dummies* regionais, por exemplo.

Uma vez testada e comprovada a existência de autocorrelação espacial nos resíduos da regressão do DID por meio de testes de autocorrelação espacial (*I* de Moran, testes *LM*, teste *CD* de Pesaran, entre outros), a pressuposição SUTVA de que o tratamento recebido por uma região não impacte o resultado de outras regiões não tratadas não pode ser garantida. Diante disso, o efeito causal do tratamento não pode ser identificado (CAMERON e TRIVEDI, 2005). Para contornar o problema da não identificação do efeito do tratamento, deve-se modelar o DID de forma a captar a interação espacial entre os agentes, flexibilizando, assim, o pressuposto SUTVA.

Nesse sentido, visando controlar a autocorrelação espacial na análise de intervenção, Dubé *et al.* (2014) modelaram o DID por meio de uma especificação

autorregressiva espacial da variável dependente por meio do modelo espacial do tipo SAR. Baseado em uma especificação SAR sobre a equação de preço hedônico, o novo estimador proposto pelos autores foi chamado de estimador de diferenças em diferenças espacial (SDID). Os autores mostraram que o estimador SDID possui vantagens em relação ao DID, a saber: i) permite testar a existência de *spillover* espacial por meio de um teste t padrão sobre o parâmetro que indica a autocorrelação espacial, ao invés de impor a pressuposição de que o efeito espacial seja inexistente e, ii) caso a autocorrelação espacial exista, pode-se decompor o efeito marginal do tratamento em efeitos total, direto e indireto.

Chagas *et al.* (2014) propuseram uma extensão do SDID ao modelar a defasagem espacial do tratamento de modo a capturar o efeito indireto do tratamento sobre todas as regiões, tratadas e não tratadas, condicionada a um vizinho tratado. Os autores aplicaram a extensão do SDID proposta para verificar o impacto da produção da cana-de-açúcar na região tratada sobre as hospitalizações em ambas as regiões, tratadas e não tratadas.

Apesar de Dubé *et al.* (2014) e Chagas *et al.* (2014) explicitarem de forma compreensível a modelagem da autocorrelação espacial em um modelo DID, Almeida (2014) alerta para o fato de que a modelagem da dependência espacial somente por meio da defasagem espacial da variável dependente e/ou por meio da defasagem espacial da variável indicadora do tratamento no DID pode não ser suficiente para tratar toda a autocorrelação espacial contida no fenômeno a ser estudado. Dessa forma, a incorporação da interação espacial entre os agentes considerará todas as possíveis formas existentes para a captação da autocorrelação espacial.

Assim, considerando o grupo de controle como um bom contrafactual e a flexibilização da premissa SUTVA na presença de dependência espacial, Almeida (2014) demonstra que uma especificação geral do SDID capaz de modelar para todas as formas existentes de autocorrelação espacial pode ser feita por meio de uma extensão da equação (3.4) do DID, como segue:

$$y_{ij}^t = \alpha + \beta_1 D^t + \beta_2 D_j + \phi D_j^t + \rho \mathbf{W} y_{ij}^t + \delta \mathbf{W} D_j^t + \mathbf{X}_{ij}^t \boldsymbol{\theta} + \mathbf{W} \mathbf{X}_{ij}^t \boldsymbol{\tau} + \xi_{ij}^t \quad (3.18a)$$

$$\xi_{ij}^t = \lambda \mathbf{W} \xi_{ij}^t + \varepsilon_{ij}^t \quad (3.18b)$$

Em que a variável $\mathbf{W}y$ denota os efeitos de interação endógena entre as variáveis dependentes; $\mathbf{W}D_j^t$ representa o efeito indireto do tratamento sobre os grupos de

tratamento e de controle; \mathbf{WX} são os efeitos de interação exógena entre as variáveis exógenas e $\mathbf{W}\xi$ referem-se aos efeitos de interação entre os termos de distúrbio das diferentes unidades espaciais; ρ é o parâmetro de defasagem espacial e λ denota o parâmetro de erro autorregressivo espacial, ambos parâmetros escalares espaciais a serem estimados; δ é um parâmetro e $\boldsymbol{\tau}$ um vetor de coeficientes, sendo que ambos são coeficientes a serem estimados referentes aos transbordamentos espaciais das variáveis exógenas; \mathbf{W} é a matriz de ponderação espacial de dimensão $N \times N$ que define o arranjo espacial das observações, sendo \mathbf{W} , por pressuposição, constante ao longo do tempo. O restante da notação permanece a mesma da equação (3.4).

O modelo que incorpora a interação espacial endógena ($\mathbf{W}y$), o efeito indireto do tratamento ($\mathbf{W}D_j^t$), a interação espacial exógena (\mathbf{WX}), e a correlação dos erros aleatórios ($\lambda\mathbf{W}$) em um só modelo, tal como descrito pela equação (3.18), é conhecido na literatura econométrico-espacial como modelo de Manski. Segundo Elhorst (2010, p. 14), em harmonia com Manski (1993), não há obstáculos técnicos para estimar o modelo descrito pela equação (3.18). Entretanto, o autor alerta para o fato de que a inferência sobre os parâmetros estimados não pode ser realizada, pois os efeitos endógenos e exógenos não podem ser distinguidos um do outro devido ao efeito de reflexão (MANSKI, 1993). Como discutido em Manski (1993), deve-se incorporar alguma restrição sobre o(s) efeito(s) espacial(is) para que os coeficientes sejam corretamente identificados.

A partir das equações (3.18a) e (3.18b) é possível obter o efeito médio do tratamento sobre o grupo tratado (ATET), bem como o efeito médio do tratamento (ATE). Analogamente ao procedimento algébrico realizado anteriormente para o DID, o ATET do SDID pode ser obtido por meio da diferença dos períodos pós e pré-intervenção, considerando os indivíduos pertencentes ao grupo de tratamento de tal forma que as equações (3.18a) e (3.18b) tornam-se:

$$\begin{aligned} (y_{i1}^1 - y_{i1}^0) &= (\alpha - \alpha) + (\beta_1 D^1 - \beta_1 D^0) + (\beta_2 D_1 - \beta_2 D_1) + (\emptyset D_1^1 - \emptyset D_1^0) + \\ &(\rho \mathbf{W}y_{i1}^1 - \rho \mathbf{W}y_{i1}^0) + (\delta \mathbf{W}D_1^1 - \delta \mathbf{W}D_1^0) + (\mathbf{X}_{i1}^1 \boldsymbol{\theta} - \mathbf{X}_{i1}^0 \boldsymbol{\theta}) + (\mathbf{W}\mathbf{X}_{i1}^1 \boldsymbol{\tau} - \mathbf{W}\mathbf{X}_{i1}^0 \boldsymbol{\tau}) + \\ &(\xi_{i1}^1 - \xi_{i1}^0) \end{aligned} \quad (3.19a)$$

$$(\xi_{i1}^1 - \xi_{i1}^0) = (\lambda \mathbf{W} \xi_{i1}^1 - \lambda \mathbf{W} \xi_{i1}^0) + (\varepsilon_{i1}^1 - \varepsilon_{i1}^0) \quad (3.19b)$$

Rearranjando os termos, o ATET pode ser estimado por meio das equações abaixo:

$$y_1 = \beta_1 + \phi D_1^1 + \rho W y_1 + \delta W D_1^1 + \Delta X_1 \theta + W \Delta X_1 \tau + v_1 \quad (3.20a)$$

$$v_1 = \lambda W v_1 + \epsilon_1 \quad (3.20b)$$

Em que, por suposição, ϵ_{i1} é um termo de erro aleatório.

Já a especificação geral para a obtenção do ATE no método SDID é definida por meio das primeiras diferenças entre os períodos pós e pré-intervenção entre os indivíduos elegíveis ao tratamento da seguinte forma:

$$(y_i^1 - y_i^0) = (\alpha - \alpha) + (\beta_1 D^1 - \beta_1 D^0) + (\beta_2 D_j - \beta_2 D_j) + (\phi D_j^1 - \phi D_j^0) + (\rho W y_i^1 - \rho W y_i^0) + (\delta W D_j^1 - \delta W D_j^0) + (X_i^1 \theta - X_i^0 \theta) + (W X_i^1 \tau - W X_i^0 \tau) + (\xi_i^1 - \xi_i^0) \quad (3.21a)$$

$$(\xi_i^1 - \xi_i^0) = (\lambda W \xi_i^1 - \lambda W \xi_i^0) + (\epsilon_i^1 - \epsilon_i^0) \quad (3.21b)$$

Rearranjando as equações (3.21a) e (3.21b), o ATE pode ser obtido por intermédio da estimação do modelo abaixo:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \rho W \Delta y_i + \delta W D_j^1 + \Delta X_i \theta + W \Delta X_i \tau + v_i \quad (3.22a)$$

$$v_i = \lambda W v_i + \epsilon_i \quad (3.22b)$$

Considerando a ressalva feita por Manski (1993) acerca da necessidade de imposição de restrição sobre os efeitos espaciais para a correta identificação dos parâmetros, vale destacar que o modelo DID convencional pode ser visto como um caso particular do modelo SDID. Para isto, basta assumir que os parâmetros espaciais da especificação geral do modelo SDID [equações (3.22a) e (3.22b)] sejam todos iguais a zero ($\rho = \lambda = \delta = 0$ e $\tau = 0$). Impondo outras restrições sobre os parâmetros espaciais, o ATET e o ATE podem ser obtidos para os diferentes modelos SDID detalhados a seguir.

3.2.1 Modelo SDID-SAR

O processo gerador dos dados do tipo SDID-SAR pode ser obtido a partir das equações (3.22a) e (3.22b), impondo restrições de que o parâmetro espacial relativo aos transbordamentos espaciais, bem como o parâmetro de erro autorregressivo espacial sejam iguais a zero ($\tau = 0$ e $\delta = \lambda = 0$). Sendo assim, as referidas equações tornam-se:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \rho W \Delta y_i + \Delta X_i \theta + \epsilon_i \quad (3.23)$$

Vale salientar que se o verdadeiro modelo a ser estimado for do tipo SDID-SAR, porém a defasagem espacial ($W \Delta y_i$) não for incluída na regressão e, assim, for estimado o modelo DID convencional (equação 3.17), incorre-se no problema de viés de variável relevante omitida. E, mesmo trabalhando com grandes amostras, o estimador MQO é inconsistente neste caso.

Ainda que a especificação do modelo seja feita de forma correta, a estimação do SDID-SAR por meio do método de MQO não é mais adequada, uma vez que tal estimador não é eficiente e consistente na presença de $W \Delta y_i$ pelo fato de a mesma tratar-se de uma variável endógena. Diante disso, é necessário adotar um estimador consistente e eficiente que considere o viés de simultaneidade causado pela inclusão da defasagem espacial.

O viés de simultaneidade é causado pela natureza multidirecional da dependência espacial, pois a presença da defasagem espacial da variável dependente ($W \Delta y_i$) é equivalente à introdução de uma variável endógena num sistema de equações simultâneas, ou seja, está correlacionada com o termo de erro aleatório (ϵ_i). Empiricamente, a endogeneidade da variável $W \Delta y_i$ pode ser testada por intermédio do teste de endogeneidade de Hausman (1978).

Diante do exposto acerca da endogeneidade de $W \Delta y_i$, um método adequado para estimar o SDID-SAR, sob a pressuposição de homocedasticidade e da normalidade dos resíduos, pode ser o método de máxima-verossimilhança (MV) (ANSELIN, 1988; ANSELIN E BERA, 1998). Contudo, a adoção deste método de estimação envolve um desafio computacional, pois devido à simultaneidade da interação espacial, aparece, na função de log-verossimilhança, o jacobiano da transformação que assume a forma de uma matriz plena, tal como $\ln[\det(\mathbf{I}_n - \rho \mathbf{W})]$, sendo \mathbf{I}_n uma matriz identidade de dimensão

$N \times N$. De acordo com Bell e Bockstael (2000), dentro da mecânica do método da máxima verossimilhança, o jacobiano é uma matriz quadrada com dimensão da amostra ($N \times N$) cujo determinante deve ser repetidamente avaliado na busca pela estimativa do coeficiente espacial (ρ). Essa característica eleva o peso computacional na etapa de estimação desses modelos, principalmente quando se trabalha com grandes amostras.

Em decorrência dessa dificuldade, Ord (1975) propôs uma simplificação para o termo do jacobiano na avaliação da função de log-verossimilhança, sugerindo substituir $\ln[\det(\mathbf{I}_n - \rho\mathbf{W})]$ por $\ln[\sum(\mathbf{I}_n - \rho\boldsymbol{\omega}_i)]$, sendo $\boldsymbol{\omega}_i$ os autovalores da matriz de ponderação espacial \mathbf{W} . Isso facilita porque o que se exige agora é computar uma única vez os autovalores, e não repetidamente. Contudo, o problema dessa solução é que a precisão dos autovalores computados decresce à medida que a dimensão da matriz \mathbf{W} aumenta (KELEJIAN e PRUCHA, 1999).

Para contornar este eventual problema computacional da estimação por MV, pode-se, alternativamente, estimar o modelo usando o método de mínimos quadrados em dois estágios (MQ2E). A adoção desse método envolve a instrumentalização da variável endógena ($\mathbf{W}\Delta y_i$). Conforme descrito em Wooldridge (2010), um bom instrumento requer que a variável Q utilizada como instrumento seja altamente correlacionada com a variável a ser instrumentalizada [$cov(Q_i, \mathbf{W}\Delta y_i) \neq 0$] e, concomitantemente, não correlacionada com o termo de erro [$cov(Q_i, \epsilon_i) = 0$].

Kelejian e Robinson (1992) e Kelejian e Prucha (1998) sugerem as defasagens espaciais das variáveis explicativas exógenas (por exemplo, $\mathbf{W}\Delta\mathbf{X}$ ou $\mathbf{W}^2\Delta\mathbf{X}$) como possíveis instrumentos para $\mathbf{W}\Delta y_i$. Se as covariadas forem realmente exógenas, as variáveis explicativas defasadas espacialmente possuem as características necessárias para serem consideradas bons instrumentos para instrumentalizar $\mathbf{W}\Delta y_i$. De qualquer modo, a validade dos instrumentos é uma propriedade empírica que pode ser testada, a *posteriori*, por meio do teste SARG proposto por Sargan (1964), por exemplo.

No caso da validação dos instrumentos, o método de MQ2E estima de forma consistente e eficiente o SDID-SAR especificado na equação (3.23). Para tanto, considere a equação (3.23) reespecificada como segue:

$$\Delta y_i = \beta_1 + \mathbf{Z}_i\boldsymbol{\pi} + \epsilon_i \quad (3.24)$$

Onde $\mathbf{Z}_i = (D_j^1, \mathbf{W}\Delta y_i, \Delta \mathbf{X}_i)$, $\boldsymbol{\pi} = (\emptyset, \rho, \boldsymbol{\theta})$ e ϵ_i é um termo de erro aleatório com média zero e variância constante, σ^2 . A partir da equação (3.24), pode-se derivar o estimador MQ2E para $\boldsymbol{\pi}$ da seguinte forma:

$$\hat{\boldsymbol{\pi}}^{MQ2E} = (\mathbf{Z}'\mathbf{P}_Q\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{P}_Q'\Delta y \quad (3.25)$$

Onde $\mathbf{P}_Q = [\mathbf{Q}(\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}']$ e \mathbf{Q} é o conjunto de instrumentos para instrumentalizar a variável endógena $\mathbf{W}\Delta y_i$, sendo $\mathbf{Q} = (\Delta \mathbf{X}, \mathbf{W}\Delta \mathbf{X})$.

Sob a suposição de homocedasticidade do vetor de erros, a variância do vetor de estimadores de MQ2E será dada por:

$$Var(\hat{\boldsymbol{\pi}}^{MQ2E}) = \hat{\sigma}^2(\mathbf{Z}'\mathbf{P}_Q\mathbf{Z})^{-1} \quad (3.26)$$

Contudo, nem sempre a premissa de homocedasticidade pode ser garantida. Na presença de heterocedasticidade, o estimador de $\boldsymbol{\pi}$ por MV pode ser inconsistente (LEE, 2004; KELEJIAN e PRUCHA, 2007; ARRAIZ *et al.*, 2010; KELEJIAN e PRUCHA, 2010), ao passo que a distribuição assintótica para o estimador de $\boldsymbol{\pi}$ por MQ2E de Kelejian e Prucha (1998) perde eficiência. Portanto, desconsiderar o problema da heterocedasticidade quando já existe, simultaneamente, o problema da dependência espacial torna as estimativas inconsistentes, uma vez que o estimador é ineficiente sob heterocedasticidade enquanto a dependência espacial pode viesar as estimativas.

Diante disso, Kelejian e Prucha (2007) desenvolveram um estimador não paramétrico, consistente, eficiente e robusto para a matriz de variância-covariância na presença de heterocedasticidade e autocorrelação espacial, denominado SHAC, para estimar um modelo representado pela equação (3.27). Tendo como referência o modelo SDID-SAR (equação 3.24), é preciso reescrever o termo de erro de tal modo que possa incluir a possibilidade da presença de heterocedasticidade:

$$\Delta y_i = \beta_1 + \mathbf{Z}_i\boldsymbol{\pi} + \xi_i \quad (3.27a)$$

$$\xi_i = \mathbf{R}\zeta \quad (3.27b)$$

$$\zeta \sim (0, \Omega) \quad (3.27c)$$

Em que ζ é um distúrbio aleatório com média 0 e variância Ω . Na equação (3.27b), tem-se \mathbf{R} , uma matriz não estocástica de dimensão $N \times N$ que pode assumir várias formas de heterocedasticidade e autocorrelação espacial simultaneamente. O restante da notação permanece a mesma da equação (3.24). A diferença aqui é que não se tenta estimar parâmetros para a autocorrelação espacial do termo de erro, como, por exemplo, estimar λ .

Dada a endogeneidade de $\mathbf{W}\Delta y_i$, o estimador de MQ2E do modelo SDID-SAR heterocedástico, tal como no modelo SDID-SAR homocedástico, é dado por:

$$\hat{\pi}^{MQ2E} = (\mathbf{Z}'\mathbf{P}_Q\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{P}_Q'\Delta y \quad (3.25)$$

Como a premissa de homocedasticidade é violada, a matriz de variância do estimador de MQ2E passa a ser a seguinte:

$$Var(\hat{\pi}^{MQ2E}) = [\mathbf{Z}'\mathbf{Q}(\mathbf{Q}'\mathbf{\Omega}\mathbf{Q})^{-1}\mathbf{Q}'\mathbf{Z}]^{-1} \quad (3.28)$$

A existência simultânea de heterocedasticidade e de autocorrelação espacial do termo de erro manifesta-se na forma da matriz $\mathbf{Q}'\mathbf{\Omega}\mathbf{Q}$. Foi desenvolvido um estimador consistente e robusto para esta matriz \mathbf{R} a fim de acomodar conjuntamente tanto a heterocedasticidade quanto a autocorrelação espacial do termo de erro (KELEJIAN e PRUCHA, 2007). Para isso, adota-se o estimador *kernel* não paramétrico para esta matriz.

Conforme alertado por Almeida (2014), convém observar que um elemento típico da variância das estimativas de MQ2E do referido modelo espacial é composto por três componentes, a saber, os elementos da matriz \mathbf{Q} , os resíduos de MQ2E das regiões i e j e a função *kernel*, que depende da distância entre as regiões (d_{ij}) e de uma distância de corte (\bar{d}). A especificação desta função *kernel* é importante a fim de fazer com que a autocorrelação espacial contida no termo de erro seja decrescente à medida que as regiões i e j estão mais distantes, formando um raio de influência para cada região¹².

Com a especificação SDID-SAR, o cálculo do efeito marginal do tratamento deve considerar o efeito de transbordamento espacial da variável dependente pela forma da matriz \mathbf{W} , bem como a magnitude do coeficiente de defasagem espacial (ρ). O efeito

¹² Existem várias especificações da função *kernel* conhecidas na literatura para formar este raio de influência, tais como uniforme, triangular, Epanechnikov, quártico, gaussiano, entre outros.

marginal expressa o impacto de uma mudança em uma variável explicativa sobre a variável dependente, assim como o efeito retroalimentador expresso pelo efeito de uma mudança em uma dada variável dependente, Δy_i , sobre as variáveis dependentes das regiões vizinhas, Δy_j , e assim por diante. Em termos formais, a partir das equações (3.24) e (3.27a), o efeito marginal pode ser expresso da seguinte forma:

$$\frac{\partial \Delta y_i}{\partial z_i} = (\mathbf{I}_n - \mathbf{W}\rho)^{-1} \boldsymbol{\pi} \quad (3.29)$$

Onde $(\mathbf{I}_n - \mathbf{W}\rho)^{-1}$ é conhecido como o multiplicador espacial. Caso \mathbf{W} seja normalizada na linha, o multiplicador espacial passa a ser $(\mathbf{I}_n - \rho)^{-1}$.

De acordo com LeSage e Pace (2009), o efeito marginal pode ser decomposto em três diferentes componentes, quais sejam: i) efeito marginal direto (equação 3.30), mensurado pela média dos elementos que aparecem na diagonal principal de $(\mathbf{I}_n - \mathbf{W}\rho)^{-1}$; ii) efeito marginal total (equação 3.31), mensurado pela média da soma das linhas ou pela média da soma das colunas da matriz $(\mathbf{I}_n - \mathbf{W}\rho)^{-1}$ e, iii) efeito marginal indireto (equação 3.32), obtido pela diferença entre o efeito marginal total e o efeito marginal direto. Formalmente, a decomposição do efeito marginal é a seguinte:

$$\overline{EM}_{direto} = N_T^{-1} \text{traço}[(\mathbf{I}_n - \mathbf{W}\rho)^{-1}] \mathbf{I}_n \boldsymbol{\pi}_k \quad (3.30)$$

$$\overline{EM}_{total} = N_T^{-1} (\mathbf{I}_n - \mathbf{W}\rho)^{-1} \mathbf{I}_n \boldsymbol{\pi}_k \quad (3.31)$$

$$\overline{EM}_{indireto} = \overline{EM}_{total} - \overline{EM}_{direto} \quad (3.32)$$

Diante disso, para computar o efeito médio do tratamento, considere o modelo SDID-SAR em sua forma reduzida, representado pela equação (3.23):

$$\Delta y_i = \beta_1 + \phi D_j^1 + \rho \mathbf{W} \Delta y_i + \Delta \mathbf{X}_i \boldsymbol{\theta} + \epsilon_i \quad (3.23)$$

Rearranjando algebricamente, obtém-se:

$$\Delta y_i = (\mathbf{I} - \mathbf{W}\rho)^{-1} \beta_1 + (\mathbf{I} - \mathbf{W}\rho)^{-1} \phi D_j^1 + \Delta \mathbf{X}_i (\mathbf{I} - \mathbf{W}\rho)^{-1} \boldsymbol{\theta} + (\mathbf{I} - \mathbf{W}\rho)^{-1} \epsilon_i \quad (3.33)$$

Assim, o efeito médio do tratamento em um modelo SDID-SAR é dado por $(\mathbf{I} - \mathbf{W}\rho)^{-1}\phi$ e não apenas por ϕ , como no modelo DID convencional. O efeito médio do tratamento considera, portanto, as interações espaciais entre todas as regiões, influenciando no impacto causal total do tratamento (LeSAGE e PACE, 2009). Cabe ressaltar que a decomposição do efeito médio do tratamento mostra-se interessante, uma vez que revelará se a interação espacial entre os agentes está suavizando ou potencializando o impacto da intervenção. Isso constitui uma informação útil para os *policy makers* no que se refere ao entendimento dos resultados efetivos das políticas públicas implementadas, ante aos seus resultados esperados, de tal sorte a subsidiar as decisões governamentais.

3.2.2 Modelo SDID-SEM

Um processo gerador dos dados do tipo SDID-SEM pode ser obtido, impondo a restrição de que os parâmetros espaciais ρ , δ e τ sejam iguais a zero. Assim a especificação geral do SDID, representada pelas equações (3.22a) e (3.22b), tomam a seguinte forma:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \Delta \mathbf{X}_i \boldsymbol{\theta} + v_i \quad (3.34a)$$

$$v_i = \lambda \mathbf{W} v_i + \epsilon_i \quad (3.34b)$$

Vale apontar que se o verdadeiro modelo a ser estimado for do tipo SDID-SEM, mas por algum fortuito for estimado um DID convencional por MQO, isto acarretaria estimativas ineficientes, uma vez que a variância do estimador de MQO seria viesada. Nesse sentido, deve-se buscar estimadores que estimem o modelo gerador dos dados do tipo SDID-SEM de forma consistente e eficiente.

Similar ao SDID-SAR, no SDID-SEM homocedástico, há o aparecimento do determinante do jacobiano da transformação do tipo $\det(\mathbf{I}_n - \lambda \mathbf{W})$ na função de log-verossimilhança, o que torna a estimação por meio do método de MQO inapropriada. Diante disso, fazendo a simplificação proposta por Ord (1975), é possível estimar

eficientemente o modelo SDID-SEM por MV. Todavia, para grandes amostras, a estimação do modelo SDID-SEM por MV pode não ser exequível em razão do ônus computacional de tal modo que a estimação por MV pode ser interrompida (KELEJIAN e PRUCHA, 1999). Isso ocorre porque, mesmo com a simplificação proposta por Ord (1975), também no âmbito do modelo SDID-SEM, há a necessidade de se computarem autovalores da matriz \mathbf{W} a fim de avaliar o jacobiano que aparece na função de log-verossimilhança, além de inverter matrizes do tipo $(\mathbf{I}_n - \lambda\mathbf{W})$.

De forma alternativa ao método MV, Kelejian e Prucha (1999) sugerem um procedimento que segue a mecânica do Método dos Momentos Generalizados (GMM) para se estimar o modelo SDID-SEM. Para tanto, é montado um sistema de equações que representam as condições de momentos e suas contrapartes amostrais para estimar consistentemente o parâmetro espacial λ que acompanha a defasagem espacial dos resíduos ($\mathbf{W}v_i$).

O procedimento de estimação proposto por Kelejian e Prucha (1999) envolve três estágios. No primeiro estágio, obtêm-se os resíduos do modelo DID convencional estimado por MQO. No segundo estágio, esses resíduos e suas respectivas defasagens espaciais são colocados no sistema de condições de momento para estimar consistentemente λ . De posse da estimativa de λ , as variáveis são filtradas de acordo com o procedimento *à la* Cochrane-Orcutt espacial:

$$\Delta y_i^* = \Delta y_i - \lambda \mathbf{W} \Delta y_i \quad (3.35a)$$

$$\Delta X_i^* = \Delta X_i - \lambda \mathbf{W} \Delta X_i \quad (3.35b)$$

Cabe salientar que o parâmetro espacial λ revela-se um parâmetro de distúrbio utilizado, apenas, para filtrar as variáveis do modelo de regressão e, conseqüentemente, o erro-padrão referente ao parâmetro espacial não é computado.

No terceiro estágio, estima-se a regressão com as variáveis filtradas espacialmente para se obter as estimativas consistentes e eficientes de $\boldsymbol{\theta}$ e $\boldsymbol{\phi}$. Após realizado o procedimento de Cochrane-Orcutt espacial, as equações (3.34a) e (3.34b) podem ser reespecificadas da seguinte forma:

$$\Delta y_i^* = \beta_1 + \mathbf{Z}_i^* \boldsymbol{\pi} + \epsilon_i \quad (3.36)$$

Onde $\mathbf{Z}_i^* = (D_j^1, \Delta \mathbf{X}_i^*)$ é uma matriz que contém as variáveis exógenas filtradas espacialmente.

O estimador para $\boldsymbol{\pi}$ pode, então, ser derivado da seguinte forma:

$$\hat{\boldsymbol{\pi}}^{MQO} = (\mathbf{Z}_i^{*\prime} \mathbf{Z}_i^*)^{-1} \mathbf{Z}_i^{*\prime} \Delta y_i^* \quad (3.37)$$

Contudo, a pressuposição de homocedasticidade nem sempre pode ser garantida. No caso em que o modelo SDID-SEM apresentar heterocedasticidade, o estimador definido acima não é mais eficiente. Diante disso, é preciso encontrar um estimador que modele a autocorrelação espacial e a heterocedasticidade, concomitantemente. Para contornar esse problema, Kelejian e Prucha (2010) propuseram um estimador para o modelo SDID-SEM para acomodar a heterocedasticidade desconhecida. Esse estimador é uma extensão do método de mínimos quadrados em dois estágios espaciais generalizados (GS2SLS)¹³.

Usando a especificação geral do modelo SDID-SEM como referência, expressa pela equação (3.35a) e, considerando a possibilidade de incorporar a heterocedasticidade no termo de erro, o processo gerador dos dados para um modelo SDID-SEM com erros heterocedásticos é dado por:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \Delta \mathbf{X}_i \boldsymbol{\theta} + v_i \quad (3.38a)$$

$$v_i = \lambda \mathbf{W} v_i + \zeta_i \quad (3.38b)$$

$$\zeta_i \sim Normal(0, \Omega) \quad (3.38c)$$

$$\Omega = \text{diag}(\sigma_i^2) \quad (3.38d)$$

Onde Ω é uma matriz diagonal composta pelos termos de variância que não são constantes e desconhecidas em sua especificação.

Para se estimar o modelo SDID-SEM com erros heterocedásticos de maneira consistente e eficiente, Kelejian e Prucha (2010) desenvolveram um método de estimação em três estágios similar ao procedimento empregado para estimar o SDID-SEM homocedástico proposto por Kelejian e Prucha (1999). A diferença entre os métodos de estimação para o SDID-SEM homocedástico e SDID-SEM heterocedástico recai no fato de que os resíduos e suas respectivas defasagens espaciais colocados no sistema de

¹³ GS2SLS é acrônimo de *Generalized Spatial Two Stage Least Squares*.

condições de momento para estimar consistentemente λ requer a incorporação da presença de heterocedasticidade no caso de SDID-SEM heterocedástico, conforme equação (3.38b). Para tanto, deve-se estimar a matriz de variância-covariância Ω por um método de correção da heterocedasticidade, a correção de White, por exemplo.

Analogamente ao SDID-SEM homocedástico, após realizado o procedimento de Cochrane-Orcutt espacial, as equações (3.38a) a (3.38d) são reespecificadas como segue:

$$\Delta y_i^\circ = \beta_1 + \mathbf{Z}_i^\circ \boldsymbol{\pi} + \epsilon_i \quad (3.39)$$

Onde $\mathbf{Z}_i^\circ = (D_j^1, \Delta \mathbf{X}_i^\circ)$ é uma matriz que contém as variáveis exógenas filtradas espacialmente corrigidas para a presença de heterocedasticidade dos resíduos.

O estimador consistente e eficiente para $\boldsymbol{\pi}$, controlando a existência de heterocedasticidade, é dado por:

$$\widehat{\boldsymbol{\pi}}^{MQO} = (\mathbf{Z}_i^{\circ\prime} \mathbf{Z}_i^\circ)^{-1} \mathbf{Z}_i^{\circ\prime} \Delta y_i^\circ \quad (3.40)$$

Kelejian e Prucha (2010) provam a consistência e eficiência dos coeficientes estimados, contidos no vetor $\widehat{\boldsymbol{\pi}}$, pelo método descrito acima e, ainda, mostram que o estimador possui normalidade assintótica.

O efeito médio do tratamento no SDID-SEM pode ser obtido, tomando a derivada parcial das equações (3.36) e (3.39) em relação à *dummy* D_j^1 como abaixo:

$$\frac{\partial \Delta y_i^*}{\partial D_j^1} = \emptyset \quad (3.41)$$

Portanto, o ATE para os modelos SDID-SEM é similar ao DID convencional. A diferença existente entre os referidos efeitos médios repousa no fato de que a estimação de \emptyset pelos métodos propostos por Kelejian e Prucha (1999; 2010) é eficiente, ou seja, os erros-padrão são robustos, uma vez que possuem variância mínima, ao contrário do DID convencional que, na presença de dependência espacial na forma de erro autorregressivo espacial, possui variância viesada.

3.2.3 Modelo SDID-SAC

Impondo a restrição de que os parâmetros espaciais δ e τ sejam iguais a zero, obtém-se o processo gerador dos dados do tipo diferenças em diferenças espacial de defasagem com erro autorregressivo (SDID-SAC). Com isso, a especificação geral do SDID, representada pelas equações (3.22a) e (3.22b), tomam a seguinte forma:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \rho \mathbf{W} \Delta y_i + \Delta \mathbf{X}_i \boldsymbol{\theta} + v_i \quad (3.42a)$$

$$v_i = \lambda \mathbf{W} v_i + \epsilon_i \quad (3.42b)$$

Interessante notar que o modelo SDID-SAC pode ser visto como uma fusão entre os modelos SDID-SAR e SDID-SEM. Sendo assim, este modelo possui uma defasagem espacial e um termo de erro espacialmente dependente, o que indica que há uma interação na variável dependente endógena (Δy_i) e no termo de erro (v_i) entre as regiões circunvizinhas.

No caso de se estimar um DID convencional por MQO quando deveria ser estimado o modelo do tipo SDID-SAC, isso geraria estimativas viesadas, inconsistentes e ineficientes. O problema de viés de variável relevante omitida ocorre pela ausência da defasagem espacial da variável dependente ($\mathbf{W} \Delta y_i$) da mesma forma que em um modelo SDID-SAR. Já a perda de eficiência das estimativas decorre da não inclusão do erro autorregressivo espacial ($\mathbf{W} v_i$), analogamente ao SDID-SEM. Diante do exposto, deve-se encontrar um estimador consistente e eficiente que acomode ambas as formas de dependência espacial encontradas nas equações (3.42a) e (3.42b).

A estimação do modelo SDID-SAC por MV envolve a otimização da função log-verossimilhança. Nessa função aparecem dois jacobianos, um vinculado à defasagem espacial da variável dependente e outro ligado ao termo de erro autocorrelacionado espacialmente, o que torna o processo de estimação por MV demasiadamente complexo, trazendo a possibilidade efetiva de que a estimação não opere.

Para contornar tal dificuldade computacional, Kelejian e Prucha (1998) propuseram um método de estimação capaz de acomodar tanto a defasagem espacial da variável dependente como o erro autorregressivo espacial, o GS2SLS. Este método estima de forma consistente e eficiente o SDID-SAC com a vantagem de ser

computacionalmente mais simples em relação ao MV e prescindir da hipótese da normalidade dos resíduos.

O procedimento de estimação do SDID-SAC, similar ao processo de estimação do SDID-SEM, envolve três estágios, a saber: i) deve-se estimar o modelo de defasagem espacial por MQ2E, adotando $\mathbf{W}\Delta\mathbf{X}$ e/ou $\mathbf{W}^2\Delta\mathbf{X}$ como conjunto de instrumentos para instrumentalizar a defasagem espacial da variável dependente ($\mathbf{W}\Delta y_i$); ii) usar os resíduos de MQ2E da estimação do modelo de defasagem espacial, efetuado no passo anterior, num sistema de condições de momento para obter a estimativa consistente do parâmetro espacial λ e, iii) utilizar a estimativa de λ para filtrar espacialmente as variáveis à *la Orcutt-Cochrane* e estimar a regressão com as variáveis filtradas por MQ2E.

Pós filtragem das variáveis dependente e explicativas e no caso da validação dos instrumentos, o modelo SDID-SAC, especificado nas equações (3.42a) e (3.42b), pode ser reescrito como segue:

$$\Delta y_i^* = \beta_1 + \mathbf{Z}_i^* \boldsymbol{\pi} + \epsilon_i \quad (3.43)$$

Onde $\mathbf{Z}_i^* = (D_j^1, \mathbf{W}\Delta y_i^*, \Delta\mathbf{X}_i^*)$, $\boldsymbol{\pi} = (\phi, \rho, \boldsymbol{\theta})$ e ϵ_i denota um termo de erro aleatório. A partir da equação (3.43) pode-se derivar o estimador MQ2E para $\boldsymbol{\pi}$ da seguinte forma:

$$\hat{\boldsymbol{\pi}}^{MQ2E} = (\mathbf{Z}^{*'} \mathbf{P}_Q \mathbf{Z}^*)^{-1} \mathbf{Z}^{*'} \mathbf{P}_Q \Delta y^* \quad (3.44)$$

Com o objetivo de acomodar a heterocedasticidade desconhecida no termo de erro, Kelejian e Prucha (2010) propuseram um estimador que se trata de uma extensão do método de GS2SLS. Este método visa modelar tanto a defasagem espacial da variável dependente quanto o erro autocorrelacionado espacialmente, considerando a heterocedasticidade presente no termo de erro. A partir das equações (3.38a) e (3.38b), pode-se especificar um modelo SDID-SAC com a presença de heterocedasticidade no termo de erro, como segue:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \rho \mathbf{W}\Delta y_i + \Delta\mathbf{X}_i \boldsymbol{\theta} + v_i \quad (3.45a)$$

$$v_i = \lambda \mathbf{W}v_i + \zeta_i \quad (3.45b)$$

$$\zeta_i \sim \text{Normal}(0, \Omega) \quad (3.45c)$$

$$\Omega = \text{diag}(\sigma_i^2) \quad (3.45d)$$

Cabe apontar que a notação das equações (3.45a) a (3.45d) é similar à descrita no modelo SDID-SEM heterocedástico, mas com a inclusão da defasagem espacial da variável dependente na equação (3.45a), conforme especificação de um modelo SDID-SAC.

O procedimento para estimar, de maneira consistente e eficiente, o modelo SDID-SAC heterocedástico proposto por Kelejian e Prucha (2010) é análogo ao procedimento de três estágios para a estimação do modelo SDID-SAC com erros homocedásticos desenvolvido por Kelejian e Prucha (1999).

O que diferencia os métodos de estimação é a necessidade de modelar a heterocedasticidade existente no termo de erro de modo que os resíduos e suas respectivas defasagens espaciais colocados no sistema de condições de momento estimem consistentemente o parâmetro de erro autorregressivo espacial λ . Como no modelo SDID-SEM heterocedástico, estima-se a matriz de variância-covariância Ω por um método de correção da heterocedasticidade do termo de erro, a correção de White, por exemplo, de modo a fornecer resíduos robustos para o estágio de filtragem espacial das variáveis de Cochrane-Orcutt.

Após o processo de filtragem espacial das variáveis e com vistas a facilitar a derivação do estimador do modelo SDID-SAC heterocedástico, as equações (3.45a) a (3.45d) podem ser reespecificadas como:

$$\Delta y_i^\circ = \beta_1 + \mathbf{Z}_i^\circ \boldsymbol{\pi} + \epsilon_i \quad (3.46)$$

Aqui, $\mathbf{Z}_i^\circ = (D_j^1, \mathbf{W}\Delta y_i^\circ, \Delta \mathbf{X}_i^\circ)$ é uma matriz que contém as variáveis exógenas filtradas espacialmente corrigidas para a presença de heterocedasticidade dos resíduos. Dessa forma, o estimador para $\boldsymbol{\pi}$ é dado por:

$$\hat{\boldsymbol{\pi}}^{MQ2E} = (\mathbf{Z}^\circ{}' \mathbf{P}_Q \mathbf{Z}^\circ)^{-1} \mathbf{Z}^\circ{}' \mathbf{P}_Q \Delta y^\circ \quad (3.47)$$

A demonstração da consistência e eficiência do procedimento de estimação para $\hat{\boldsymbol{\pi}}$ descrito acima bem como a garantia da normalidade assintótica do estimador é descrita de forma pormenorizada em Kelejian e Prucha (2010).

Como a especificação SDID-SAC leva em consideração também a defasagem espacial da variável dependente, o cálculo do efeito marginal do tratamento é similar ao do SDID-SAR. Dessa forma, o efeito médio é descrito como segue:

$$\frac{\partial \Delta y_i^*}{\partial D_j^1} = (\mathbf{I}_n - \mathbf{W}\rho)^{-1}\phi \quad (3.48)$$

Portanto, similarmente a um modelo SDID-SAR, o efeito marginal do tratamento é dado por $(\mathbf{I} - \mathbf{W}\rho)^{-1}\phi$ e não apenas por ϕ e, assim, pode ser decomposto em efeitos direto e indireto. A diferença está no fato de que o parâmetro ϕ foi estimado de maneira a acomodar, inclusive, a dependência espacial na forma de erro autorregressivo espacial.

3.2.4 Modelo SDID-SLX

A forma estrutural do modelo regressivo cruzado espacial (SLX) pode ser obtida impondo as restrições, nas equações (3.22a) e (3.22b) de que os parâmetros espaciais de alcance global, isto é, ρ e λ sejam iguais a zero. Dessa forma, o modelo SDID-SLX pode ser definido como segue:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \delta \mathbf{W}D_j^1 + \Delta \mathbf{X}_i \boldsymbol{\theta} + \mathbf{W} \Delta \mathbf{X}_i \boldsymbol{\tau} + \epsilon_i \quad (3.49)$$

Conforme apontado por Almeida (2014), a inclusão das defasagens espaciais do tratamento e das variáveis explicativas ($\mathbf{W}D_j^1$ e $\mathbf{W} \Delta \mathbf{X}_i$, respectivamente) no lado direito do modelo de regressão DID pode ser justificada tanto por razões teóricas quanto empíricas. Teoricamente, a expectativa da emergência de transbordamentos e externalidades espaciais, os efeitos de vizinhança, o efeito contágio, a difusão tecnológica, *copycatting*, entre outros, podem ser motivos que justifiquem a inclusão das variáveis exógenas defasadas espacialmente. Por outro lado, a fonte da presença da dependência espacial na forma de defasagem espacial das covariadas pode ser de ordem puramente empírica na forma de erro de medida das variáveis e/ou a má especificação do modelo.

Considerando que D_j^1 e ΔX_i sejam realmente exógenas, é plausível afirmar que WD_j^1 e $W\Delta X_i$ também serão exógenas. Em decorrência disso, os coeficientes espaciais δ e τ terão um efeito local, uma vez que os mesmos não afetam o sistema como um todo pelo fato de que o transbordamento das regiões vizinhas são localizados, não havendo, por conseguinte, efeito retroalimentador, como nos modelos que possuem os efeitos espaciais de alcance global ρ e λ .

Em virtude de $W\Delta X_i$ e WD_j^1 consistirem nas defasagens espaciais de ΔX_i e D_j^1 , respectivamente, é provável que exista alta correlação entre essas variáveis, o que, por sua vez, pode gerar multicolinearidade. Esse problema pode tornar as estimativas dos parâmetros da regressão instáveis, tendo em vista que o efeito marginal de uma variável explicativa específica sobre a variável dependente será dado pelo seu próprio coeficiente mais uma parcela do coeficiente da variável com a qual está altamente correlacionada.

Uma maneira de testar a presença de multicolinearidade entre as variáveis exógenas é a estimação da matriz de correlação entre essas variáveis. Caso seja verificada multicolinearidade alta entre duas variáveis explicativas e, não haja argumentos teóricos que justifiquem a inclusão das mesmas, uma das variáveis deve ser eliminada do modelo. A partir de então, há de se encontrar o método adequado para a estimação do modelo SDID-SLX.

Uma vez testada e comprovada a exogeneidade das variáveis explicativas, assim como verificada a inexistência de alta correlação entre as mesmas, o modelo SDID-SLX pode ser estimado de forma consistente e eficiente por MQO ou MV (REY e MONTOURI, 1999).

De modo a demonstrar o estimador do modelo SDID-SLX, a equação (3.49) pode ser reespecificada da seguinte maneira:

$$\Delta y_i = \beta_1 + \mathbf{Z}_i \boldsymbol{\pi} + \epsilon_i \quad (3.50)$$

Onde $\mathbf{Z}_i = (D_j^1, WD_j^1, \Delta X_i, W\Delta X_i)$ e $\boldsymbol{\pi}$ refere-se ao vetor de parâmetros que acompanha as variáveis explicativas contidas em \mathbf{Z}_i . A partir da equação acima, o estimador de MQO é dado a seguir:

$$\hat{\boldsymbol{\pi}}^{MQO} = (\mathbf{Z}_i' \mathbf{Z}_i)^{-1} \mathbf{Z}_i' \Delta y_i \quad (3.51)$$

Apesar da semelhança do estimador SDID-SLX com o estimador do DID convencional, caso seja estimado um DID não espacial quando o verdadeiro processo gerador dos dados for o SDID-SLX, incorrer-se-á no viés de variável relevante omitida.

Cabe salientar que se for verificado o problema de heterocedasticidade no termo de erro, o estimador descrito na equação (3.51) deve ser estimado com a matriz de variância-covariância de White, calculando-se os erros-padrão robustos com o intuito de resolver o problema da heterocedasticidade.

Dado que o efeito espacial contemplado no modelo SDID-SLX é de alcance local, o efeito médio do tratamento pode ser obtido a partir da derivada parcial da equação (3.49) como segue:

$$\frac{\partial \Delta y_i}{\partial D_j^1} = \phi + \delta W \quad (3.52)$$

A diferença entre os efeitos médios do tratamento do DID convencional e SDID-SLX consiste na eliminação do viés de variável relevante no último modelo em virtude da inclusão das externalidades espaciais quando tais *spillovers* espaciais forem importantes para explicar o fenômeno estudado, além do efeito indireto do tratamento sobre todas as regiões, tratadas e não tratadas.

3.2.5 Modelo SDID-SDM

De acordo com a taxonomia apresentada, o modelo Durbin espacial (SDM) contém a defasagem espacial da variável dependente ($W\Delta y_i$), bem como as defasagens espaciais das variáveis explicativas (WD_j^1 e $W\Delta X_i$), assumidas como exógenas, na sua especificação. Portanto, o modelo SDID-SDM é obtido a partir da especificação geral do SDID (equações 3.22a e 3.22b) impondo a restrição de que o parâmetro espacial λ seja igual a zero. Assim, a forma funcional do SDID-SDM é a que segue:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \rho W\Delta y_i + \delta W D_j^1 + \Delta X_i \theta + W\Delta X_i \tau + \epsilon_i \quad (3.53)$$

Como no modelo SDID-SAR, devido à endogeneidade da defasagem espacial da variável dependente, o método de MQO não é capaz de estimar de maneira consistente e eficiente tal modelo e, portanto, deve-se buscar métodos adequados para estimar o modelo SDID-SDM.

Apesar do esforço computacional previamente mencionado para a estimação de modelos que contenham a dependência espacial na forma da defasagem espacial da variável dependente ($W\Delta y_i$) em sua especificação, o método de MV, sob a hipótese de normalidade dos resíduos, pode ser aplicado para estimar consistente e eficientemente o SDID-SDM.

Todavia, um método amplamente utilizado para tratar a endogeneidade de $W\Delta y_i$, com a vantagem de prescindir da normalidade dos resíduos, é o método de MQ2E. Este método permite obter estimativas consistentes e eficientes do modelo SDID-SDM. Neste caso, deve-se buscar um conjunto de instrumentos diferente do adotado nos modelos SDID-SAR e SDID-SAC, uma vez que $W\Delta X_i$ consta na especificação do modelo SDID-SDM e, portanto, não serve como instrumento dele próprio.

Diante disso, os candidatos para instrumentos da variável endógena $W\Delta y_i$ podem ser as defasagens espaciais das defasagens espaciais das características das regiões ($W^2\Delta X_i$)¹⁴ ou a defasagem espacial e temporal da variável dependente ($W\Delta y_{i,t-1}$). Novamente, deve-se atentar para possibilidade de se ter instrumentos fracos. Ainda, cabe ressaltar que, semelhante ao modelo SDID-SLX, deve-se ter o devido cuidado para evitar o problema da multicolinearidade.

De modo a demonstrar que o método de MQ2E estima de forma consistente e eficiente o SDID-SDM, considere a equação (3.53) reespecificada como segue:

$$\Delta y_i = \beta_1 + Z_i\pi + \epsilon_i \quad (3.54)$$

Onde $Z_i = (D_j^1, W\Delta y_i, WD_j^1, \Delta X_i, W\Delta X_i)$, $\pi = (\phi, \rho, \delta, \theta, \tau)$. A partir da equação (3.24), pode-se derivar o estimador MQ2E para π da seguinte forma:

$$\hat{\pi}^{MQ2E} = (Z'P_Q Z)^{-1} Z'P_Q' \Delta y \quad (3.55)$$

¹⁴ Ou ainda defasagens espaciais de ordens superiores, tais como $W^3\Delta X_i$ ou $W^4\Delta X_i$.

A diferença entre o estimador de MQ2E do SDID-SDM (equação 3.55) e o do modelo SDID-SAR (equação 3.25) é atribuída ao conjunto de instrumentos \mathbf{Q} para instrumentalizar a variável endógena $\mathbf{W}\Delta y_i$, uma vez que $\mathbf{Q} = (\Delta \mathbf{X}, \mathbf{W}\Delta \mathbf{X})$ para o modelo SDID-SAR e para o modelo SDID-SDM $\mathbf{Q} = (\Delta \mathbf{X}, \mathbf{W}\Delta \mathbf{X}, \mathbf{W}^2\Delta \mathbf{X})$.

Na presença de heterocedasticidade, como mencionado anteriormente, os estimadores de MV são inconsistentes e, por outro lado, a distribuição assintótica para o estimador de MQ2E de Kelejian e Prucha (1998) não é apropriada. Diante disso, deve-se utilizar o estimador SHAC proposto por Kelejian e Prucha (2007), visto que este método é capaz de estimar os parâmetros do modelo SDID-SDM de maneira consistente e eficiente, uma vez que o mesmo controla, concomitantemente, a presença de heterocedasticidade no termo de erro e da endogeneidade de $\mathbf{W}\Delta y_i$.

O efeito médio do tratamento do modelo SDID-SDM é obtido de maneira similar ao efeito médio do tratamento do modelo SDID-SAR, tendo em vista que ambos os modelos possuem o efeito espacial de alcance global representado pela defasagem espacial da variável dependente ($\mathbf{W}\Delta y_i$). Com isso, o efeito médio do tratamento pode ser decomposto em efeitos direto e indireto. A diferença do efeito médio do tratamento no modelo SDID-SDM em relação ao modelo SDID-SAR é creditada à presença conjunta dos efeitos espaciais de alcance global ($\mathbf{W}\Delta y_i$), bem como de alcance local ($\mathbf{W}D_j^1$ e $\mathbf{W}\Delta \mathbf{X}_i$). Sendo assim, o efeito médio do tratamento do modelo SDID-SDM, após rearranjos algébricos da equação (3.53), é calculado como segue:

$$\frac{\partial \Delta y_i}{\partial D_j^1} = (\mathbf{I}_n - \mathbf{W}\rho)^{-1}(\emptyset + \delta \mathbf{W}) \quad (3.56)$$

3.2.6 Modelo SDID-SDEM

O verdadeiro processo gerador dos dados de um modelo de Durbin espacial de erro (SDEM) é obtido impondo a restrição, na especificação geral do SDID (equações 3.22a e 3.22b), de que o parâmetro espacial ρ seja igual a zero. Com isso, a forma funcional do SDID-SDEM passa ser a que segue:

$$\Delta y_i = \beta_1 + \phi D_j^1 + \delta W D_j^1 + \Delta X_i \theta + W \Delta X_i \tau + v_i \quad (3.57a)$$

$$v_i = \lambda W v_i + \epsilon_i \quad (3.57b)$$

O modelo SDID-SDEM incorpora componentes de transbordamento espacial de alcance local ($W D_j^1$ e $W \Delta X_i$) e de alcance global ($W v_i$).

Similarmente ao modelo SDID-SEM, caso seja estimado um DID convencional por MQO quando o verdadeiro modelo a ser estimado for do tipo SDID-SDEM, isto acarretaria estimativas ineficientes, uma vez que a variância do estimador de MQO seria viesada. Entretanto, a estimação de um modelo SDID-SDEM por MQO possui o agravante de se obter estimativas inconsistentes, em virtude da omissão de variáveis explicativas relevantes ($W D_j^1$ e $W \Delta X_i$).

Nesse sentido, para a estimação do modelo SDID-SDEM haverá o aparecimento do determinante do jacobiano da transformação do tipo $\det(I_n - \lambda W)$ na função de log-verossimilhança o que, mesmo com a simplificação proposta por Ord (1975), a estimação consistente e eficiente por MV pode não ser factível para grandes amostras. Isso pode ser decorrência do ônus em termos computacionais, para se computar os autovalores da matriz W a fim de avaliar o jacobiano que aparece na função de log-verossimilhança.

Alternativamente ao método MV, Kelejian e Prucha (1999) sugerem um procedimento que segue a mecânica do GMM para estimar o modelo SDID-SDEM. A diferença entre o procedimento de estimação para o modelo SDID-SEM e o modelo SDID-SDEM recai no estágio da filtragem das variáveis, pois o procedimento *à la* Cochrane-Orcutt passa a exigir, além da filtragem das variáveis dependente e explicativas, a filtragem das covariadas defasadas espacialmente para a adequada estimação do modelo SDID-SDEM. Dessa forma, o procedimento de filtragem espacial passa a ser o seguinte:

$$\Delta y_i^* = \Delta y_i - \lambda W \Delta y_i \quad (3.35a)$$

$$D_j^{1*} = D_j^1 - \lambda W D_j^1 \quad (3.35b)$$

$$W D_j^{1*} = W D_j^1 - \lambda W^2 D_j^1 \quad (3.35c)$$

$$\Delta X_i^* = \Delta X_i - \lambda W \Delta X_i \quad (3.35d)$$

$$W \Delta X_i^* = W \Delta X_i - \lambda W^2 \Delta X_i \quad (3.35e)$$

De posse das variáveis filtradas espacialmente, pode-se estimar os parâmetros do modelo SDID-SDEM por meio do mesmo estimador apresentado na equação (3.37) para a estimação do modelo SDID-SEM.

Com a finalidade de estimar o modelo SDID-SDEM com erros heterocedásticos, o estimador de Kelejian e Prucha (2010), descrito anteriormente, considera agora a inclusão dos transbordamentos espaciais representado por WD_j^1 e $W\Delta X_i$ na equação (3.38a). Adotando novamente o procedimento de filtragem espacial das variáveis descrito pelas equações (3.35a) a (3.35e), pode-se obter as estimativas consistentes e eficientes do modelo SDID-SDEM heterocedástico.

Da mesma forma que o efeito médio do tratamento no SDID-SEM, o efeito médio do tratamento do modelo SDID-SDEM pode ser obtido tomando a derivada parcial das equações (3.57a) e (3.57b) em relação a D_j^1 , após realizado o procedimento de filtragem espacial:

$$\frac{\partial \Delta y_i^*}{\partial D_j^1} = \phi + \delta W \quad (3.41)$$

3.2.7 Resumo dos métodos de estimação dos modelos SDID

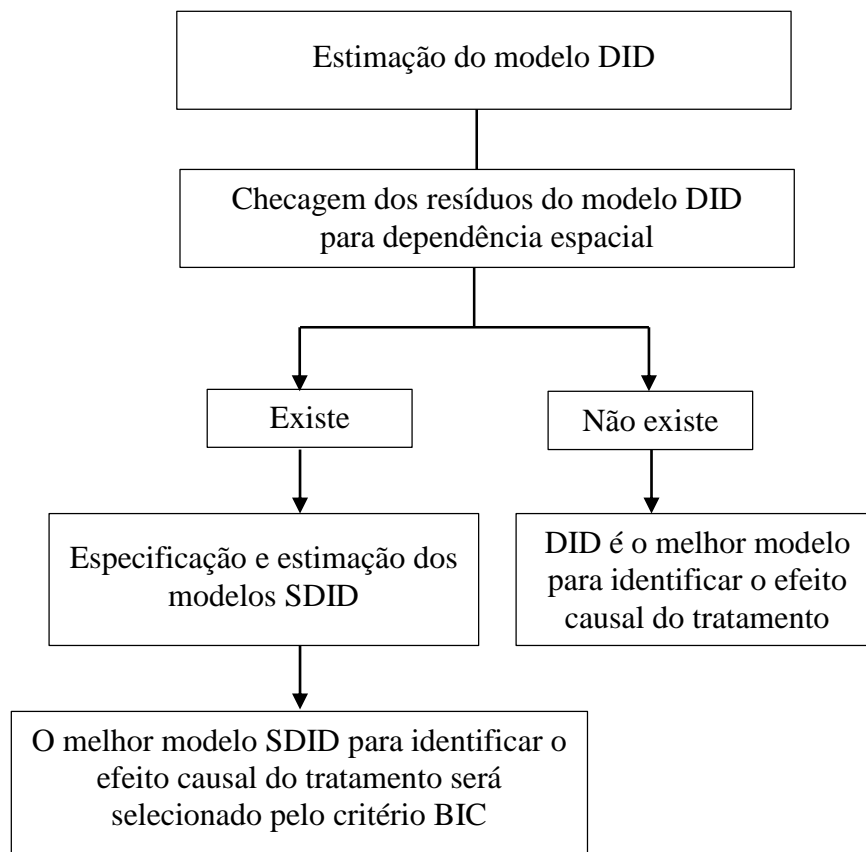
Conforme descrito anteriormente, o controle da dependência espacial, quando presente no fenômeno sob estudo, pode ser obtido por meio da reespecificação do modelo DID. Entretanto, a especificação de um modelo SDID exige a adoção de um estimador específico capaz de estimar de maneira consistente e eficiente cada modelo, de acordo com o tipo de dependência espacial observada. O Quadro 3.2 resume os principais aspectos metodológicos necessários para estimar os diversos modelos SDID de acordo com a taxonomia apresentada nas subseções anteriores. Cabe à Figura 3.1 resumir de maneira ilustrativa o procedimento de estimação e a consequente escolha do estimador que melhor identifica o efeito causal do tratamento.

Quadro 3.2- Resumo dos principais modelos de diferenças em diferenças espaciais

Modelo	Defasagem Espacial	Homocedasticidade	Consequências	Estimador Apropriado	Efeito Médio do Tratamento (ATE)
DID	Nenhuma	Sim	_____	OLS	ϕ
		Não	Ineficiência	OLS com erros-padrão robustos	ϕ
SDID-SAR	$\mathbf{W}\Delta y_i$	Sim	Inconsistência	MV ou 2SLS	$(\mathbf{I} - \mathbf{W}\rho)^{-1}\phi$
		Não	Inconsistência e ineficiência	Estimador HAC	$(\mathbf{I} - \mathbf{W}\rho)^{-1}\phi$
SDID-SEM	$\mathbf{W}v_i$	Sim	Ineficiência	MV ou GMM de Kelejian e Prucha (1999)	ϕ
		Não	Ineficiência	Método KP	ϕ
SDID-SAC	$\mathbf{W}\Delta y_i$ e $\mathbf{W}v_i$	Sim	Inconsistência e ineficiência	MV ou GS2SLS	$(\mathbf{I} - \mathbf{W}\rho)^{-1}\phi$
		Não	Inconsistência e ineficiência	Método KP	$(\mathbf{I} - \mathbf{W}\rho)^{-1}\phi$
SDID-SLX	$\mathbf{W}\Delta X_i$	Sim	Inconsistência	OLS	$\phi + \delta\mathbf{W}$
		Não	Inconsistência e Ineficiência	OLS com erros-padrão robusto	$\phi + \delta\mathbf{W}$
SDID-SDM	$\mathbf{W}\Delta y_i$ e $\mathbf{W}\Delta X_i$	Sim	Inconsistência	MV ou 2SLS	$(\mathbf{I} - \mathbf{W}\rho)^{-1}(\phi + \delta\mathbf{W})$
		Não	Inconsistência e Ineficiência	Estimador HAC	$(\mathbf{I} - \mathbf{W}\rho)^{-1}(\phi + \delta\mathbf{W})$
SDID-SEM	$\mathbf{W}v_i$ e $\mathbf{W}\Delta X_i$	Sim	Inconsistência e Ineficiência	MV ou GMM de Kelejian e Prucha (1999)	$\phi + \delta\mathbf{W}$
		Não	Inconsistência e Ineficiência	Método KP	$\phi + \delta\mathbf{W}$

Fonte: Adaptado de Almeida (2014).

Figura 3.1 - Resumo do procedimento de estimação dos modelos SDID



Fonte: Elaboração Própria.

3.3 A TEORIA ECONÔMICA DA ESCOLHA RACIONAL: DA OFERTA INDIVIDUAL À OFERTA AGREGADA DO CRIME

Cressey (1968) afirma que uma teoria para explicar o comportamento social, em particular o comportamento criminoso, deveria considerar a compreensão das motivações e do comportamento individual, bem como a epidemiologia associada, ou como tais comportamentos se distribuem e se deslocam espacial e temporalmente. Nessa perspectiva, Kelly (2000) argumenta que as mais influentes teorias ecológicas do crime podem ser resumidas em três categorias: i) a teoria econômica do crime de Becker (1968); ii) a teoria da anomia de Merton (1938) e; iii) a teoria da desorganização social de Shaw e McKay (1942).

Vale ressaltar que, apesar dessas teorias abordarem o comportamento criminoso sob prismas diferentes, Kelly (2000) argumenta que as mesmas não podem ser vistas como substitutas, mas como complementares, uma vez que tais teorias ecológicas buscam explicar as variações nos índices de criminalidade através dos diferentes incentivos, pressões e efeitos dissuasórios com que os indivíduos se deparam em diferentes ambientes. A teoria da desorganização social, por exemplo, considera as dissuasões sociais informais, enquanto a teoria da anomia concentra-se nas pressões sociais pelo sucesso econômico para o cometimento de crimes violentos. Já a teoria econômica do crime, embora formalmente capaz de abranger as demais teorias, tem como foco principal os incentivos para cometer crimes e o efeito dissuasório do sistema de justiça criminal formal.

Portanto, considerando a ressalva de Kelly (2000), o comportamento individual do criminoso pode ser compreendido no contexto da teoria econômica do crime. Cabe lembrar que os primórdios da investigação econômica do crime podem ser encontrados em Beccaria (1764) e Bentham (1789) e, posteriormente, com os estudos de Fleisher (1963, 1966) e Ehrlich (1967). Entretanto, a formalização dos modelos econômicos do crime é creditada a Becker (1968) e Ehrlich (1973). Sendo assim, o comportamento criminoso, detalhado a seguir, é baseado nos modelos de escolha racional do criminoso de Becker (1968), Ehrlich (1973) e Kelly (2000) e consiste na tentativa de evidenciar os diferentes determinantes da criminalidade, em particular, o impacto de uma política de segurança pública na oferta individual do crime e, por conseguinte, na oferta agregada do crime.

Para Becker (1968), a existência e a expansão (ou declínio) da indústria do crime, assim como outra qualquer, resultam do mercado. Nesse contexto, a premissa básica do modelo econômico do crime é que todos os indivíduos são movidos a incentivos e, assim, os criminosos se comportam racionalmente no sentido de que agem de forma a maximizar o seu bem-estar econômico (BECKER, 1968; EHRLICH, 1973).

De modo a formalizar o modelo de oferta do crime, considere N como a população de certa região e χ como a fração da população predisposta a cometer crime, sendo χ função crescente da densidade populacional d . Assuma que j refere-se a um indivíduo racional e que responde a incentivos, sendo j um indivíduo representativo de χ . A oferta individual de crime pode ser entendida, então, como um processo de maximização da utilidade esperada (UE) em que o indivíduo representativo j avalia, por um lado, os potenciais benefícios e ganhos pecuniários da ação criminosa (W_{ji}); o custo de planejamento e execução do crime (c_{ji}) e as probabilidades de detenção e aprisionamento associadas à atividade criminosa (π). Por outro lado, o custo de oportunidade de cometer crime, pode ser entendido como o salário do indivíduo j no mercado de trabalho formal (W_{jl}).

Um modelo de escolha entre atividades legais (L) e ilegais (I) pode ser formulado, portanto, dentro da estrutura da teoria econômica da escolha sob incerteza. A hipótese central dessa teoria é que, se em um dado período as duas atividades forem mutuamente exclusivas, uma atividade seria escolhida entre elas comparando a utilidade esperada associada a cada uma. Assim, o indivíduo j estaria disposto a cometer crime, caso o retorno marginal esperado em I for maior que o retorno esperado em L . Diante do exposto, os retornos esperados em I e L , são, respectivamente:

$$R_I = (1 - \pi)W_{ji} - c_{ji} - \alpha W_{jl} \quad (3.42a)$$

$$R_L = \alpha W_{jl} \quad (3.42b)$$

Onde α é a probabilidade de o indivíduo j estar empregado no mercado de trabalho formal. Para Ehrlich (1973), quaisquer atividades produzem riqueza e bem-estar ao indivíduo e estão sujeitas à incerteza. Diante disso, no caso do criminoso, ele auferir rendas pecuniárias (W_{ji}) sujeitas a um risco de condenação (π). O aumento de π reduz sua probabilidade de estar empregado no mercado formal (α) e, por conseguinte, diminui seus ganhos no mercado legal (W_{jl}). Por outro lado, o indivíduo auferir renda no mercado

legal (W_{jl}), mas está sujeito às incertezas do mercado de trabalho formal (α). Cabe apontar que alguns indivíduos com inclinação à violência podem cometer crime mesmo quando R_I for menor do que R_L .

Considerando que c_{ji} tende a zero ($c_{ji} \rightarrow 0$) pelo fato de ser o custo de planejamento e execução do crime ser relativamente baixo, Ehrlich (1973) especifica uma função comportamental da participação efetiva do indivíduo j na atividade L em determinado período - ou oferta individual do crime - controlando para os diversos determinantes como segue:

$$O_j = f(\pi, W_{ji}, W_{jl}, \alpha, \mathbf{x}_j) \quad (3.43)$$

Onde O_j é a quantidade de crimes cometida pelo indivíduo j ; \mathbf{x}_j denota todas as outras variáveis que podem afetar a frequência de crimes cometidos pelo indivíduo j , como, por exemplo, escolaridade, idade, sexo, instabilidade familiar, religião entre outras.

O modelo descrito pela equação (3.43) é de natureza microeconômica, tendo em vista que analisa os determinantes individuais da criminalidade. Considerando a premissa de que j denota um indivíduo representativo da fração da população predisposta a cometer crime (χ), a equação (3.43) pode ser agregada de modo a modelar a função comportamental de todos os indivíduos de χ em um dado período de tempo. Formalmente, tem-se:

$$\sum_{j=1}^{N\chi} O_j = O = f(\Pi, W_i, W_l, u_l, \mathbf{X}) \quad (3.44)$$

Em que O pode ser entendida como a quantidade de crimes cometida em determinada região e u_l refere-se à taxa de desemprego no mercado formal com que os indivíduos se deparam. Cabe destacar que, após a agregação da oferta individual do crime, \mathbf{X} passa a incluir também as variáveis ambientais tais como: distribuição de renda, densidade demográfica e outras.

A equação (3.44) é útil, pois permite a consecução de pesquisas a partir de uma estrutura de dados agregados regionalmente, além de ser possível demonstrar o impacto esperado de cada um dos determinantes da criminalidade sobre a quantidade de crime de certa região.

Para Becker (1968) e Ehrlich (1973), a probabilidade de detenção e aprisionamento associada à atividade criminosa (Π) representa uma medida de risco para os criminosos. Assim, Π pode ser entendido como o rigor da atividade policial e do judiciário, endossado por políticas de segurança pública, de tal modo que O tende a diminuir, quando os indivíduos se deparam com o aumento de Π ($\frac{\partial O}{\partial \Pi} < 0$). Sendo assim, políticas de segurança pública que visam à aplicação mais rigorosa da lei, como o INFOCRIM, tendem a aumentar a probabilidade de detenção e aprisionamento de modo a diminuir o retorno esperado da atividade ilegal e, conseqüentemente, reduzir a oferta de crimes.

Becker (1968) e Ehrlich (1973) mostram um efeito ambíguo da renda sobre a criminalidade. Por um lado, a maior pujança econômica aumenta o custo de oportunidade, pois torna-se possível obter maior renda no mercado formal, o que pode reduzir a criminalidade ($\frac{\partial O}{\partial w_i} < 0$). Por outro lado, o aumento da riqueza dos indivíduos leva a um maior contingente de vítimas potenciais e economicamente atrativas, gerando maiores benefícios decorrentes da atividade criminal ($\frac{\partial O}{\partial w_i} > 0$) e, por conseguinte, espera-se que haja um aumento de O .

De acordo com Kelly (2000), a densidade populacional d aumenta χ e, logo, aumentaria O - ($\frac{\partial O}{\partial d} > 0$) - de duas formas: i) aumenta a oferta das vítimas potenciais, que não conhecem o criminoso e ii) aumenta o anonimato ao reduzir a probabilidade de ser capturado (GLAESER e SACERDOTE, 1999).

No que tange à taxa de desemprego no mercado formal u_i , modelos e teorias inspiradas nos modelos econômicos de escolha racional (BECKER, 1968; ERLICH, 1973) e na teoria da anomia desenvolvida por Merton (1938)¹⁵ argumentam que o incremento na taxa de desemprego estimularia a prática criminosa ($\frac{\partial O}{\partial u_i} > 0$), visto que reduziria o custo de oportunidade dos indivíduos de cometerem atos criminosos.

Por fim, em virtude de X corresponder a uma matriz de diferentes determinantes - individuais e ambientais - que afetariam O de formas distintas, o impacto esperado de

¹⁵ Segundo a teoria da anomia de Merton (1938), a motivação para a delinquência decorreria da impossibilidade de o indivíduo atingir metas desejadas por ele, como, por exemplo, o sucesso econômico. Sendo assim, quando o indivíduo não encontra oportunidades para atingir seus objetivos no mercado legal (por exemplo, encontra-se desempregado) recorre ao mercado ilegal para alcançar seus desejos.

cada um dos determinantes abordados por essa tese serão explicitados na próxima seção (3.4).

3.4 ESTRATÉGIA EMPÍRICA E DESCRIÇÃO DAS VARIÁVEIS

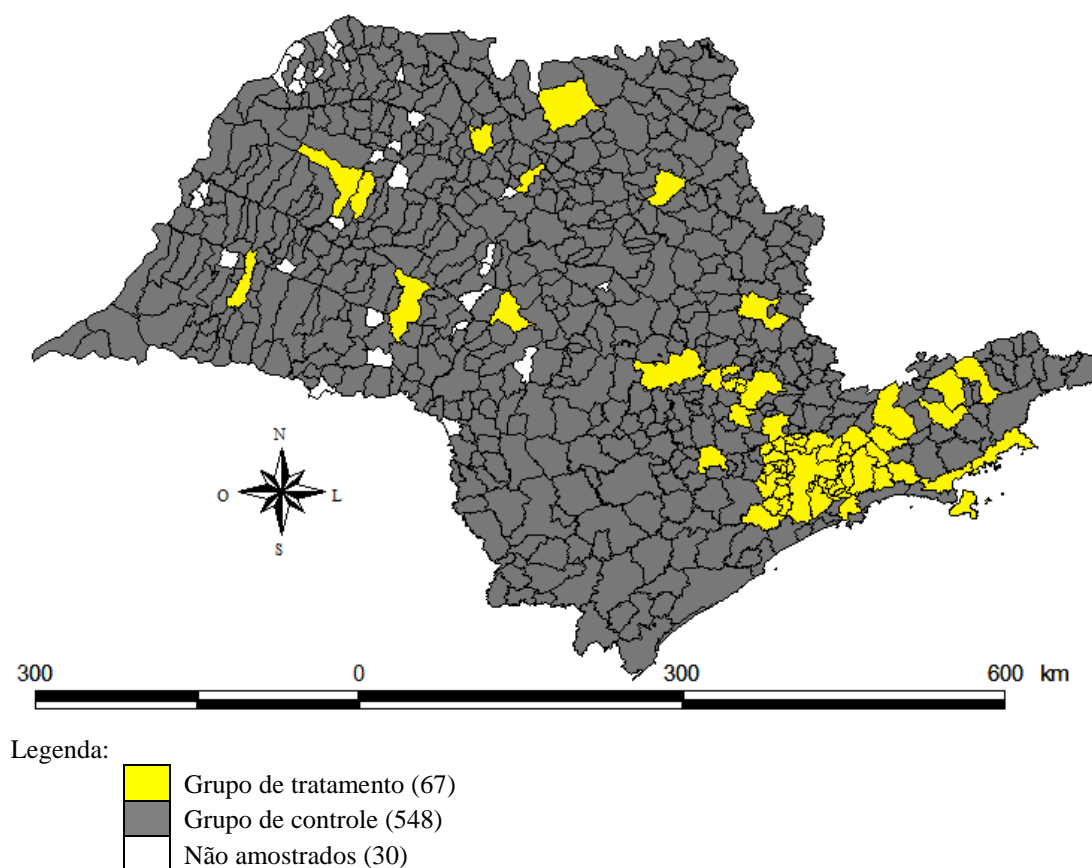
A estratégia empírica adotada por essa tese para mensurar o impacto do INFOCRIM sobre a criminalidade violenta nos municípios paulistas consistirá na aplicação dos métodos de avaliação de impacto de tratamento descritos nas seções 3.1 e 3.2. Para a implementação da metodologia de DID e SDID é necessária a definição dos grupos de tratamento e de controle. O grupo de tratamento refere-se aos 39 municípios da RMSP¹⁶ e as demais 28 cidades¹⁷ onde o INFOCRIM foi implementado até o ano de 2010, contabilizando 67 municípios.

Por sua vez, o grupo de controle foi definido como os demais municípios paulistas, cujos dados estavam disponíveis, que não foram contemplados com o Programa, totalizando 548 municípios. Os grupos de tratamento e de controle totalizam 615 municípios, contemplando 95,35% do total dos municípios do estado de São Paulo e podem ser visualizados na Figura 3.2.

¹⁶ Os 39 municípios da RMSP são: Arujá, Barueri, Biritiba-Mirim, Caieiras, Cajamar, Carapicuíba, Cotia, Diadema, Embu, Embu-Guaçu, Ferraz de Vasconcelos, Francisco Morato, Franco da Rocha, Guararema, Guarulhos, Itapeverica da Serra, Itapevi, Itaquaquetuba, Jandira, Juquitiba, Mairiporã, Mauá, Mogi das Cruzes, Osasco, Pirapora do Bom Jesus, Poá, Ribeirão Pires, Rio Grande da Serra, Salesópolis, Santa Isabel, Santana de Parnaíba, Santo André, São Bernardo do Campo, São Caetano do Sul, São Lourenço da Serra, São Paulo, Suzano, Taboão da Serra, Vargem Grande Paulista.

¹⁷ As demais vinte e oito cidades onde o INFOCRIM foi implementado são: Americana, Araçatuba, Barretos, Bauru, Birigui, Campinas, Caraguatatuba, Catanduva, Guaratinguetá, Hortolândia, Ilhabela, Indaiatuba, Jacareí, Jundiaí, Marília, Mogi Guaçu, Pindamonhangaba, Piracicaba, Presidente Prudente, Ribeirão Preto, Santa Bárbara d'Oeste, Santos, São José do Rio Preto, São José dos Campos, Sorocaba, Sumaré, Taubaté e Ubatuba.

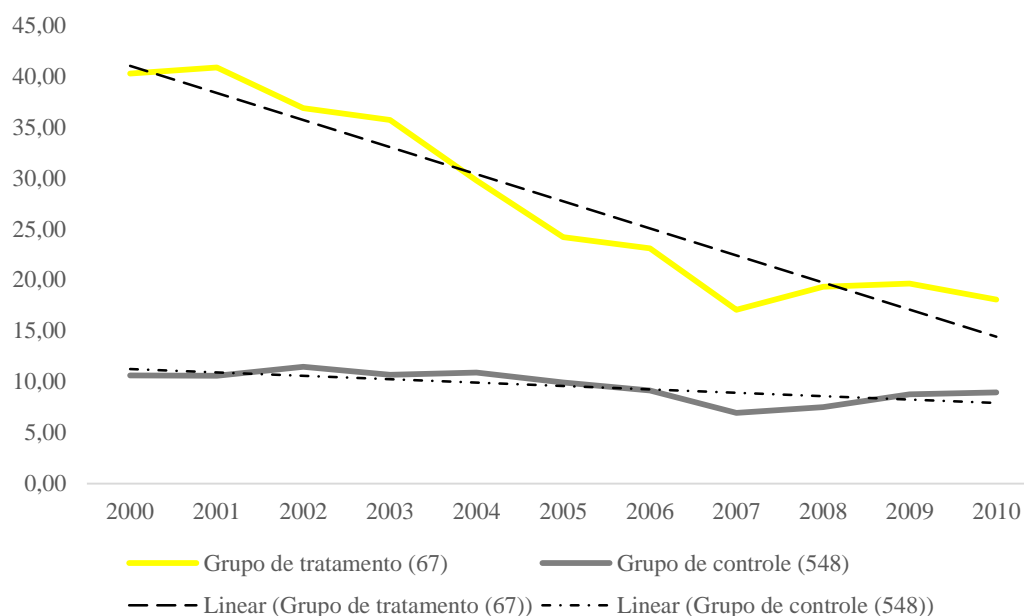
Figura 3.2 - Grupos de tratamento e de controle para a implementação do DID e SDID.



Fonte: Elaboração Própria.

O grupo de controle pode ser considerado um bom contrafactual, uma vez que as taxas de homicídios por cem mil habitantes dos grupos de tratamento e de controle exibem uma trajetória paralela declinante no início do período de análise, conforme pode ser observado por intermédio da Figura 3.3. Após a implementação do INFOCRIM, pode-se perceber que o grupo de tratamento passa a exibir uma queda abrupta e sistemática nas taxas de homicídios, enquanto o grupo de controle mantém a trajetória das taxas de crimes letais relativamente estável. Esse fato indica a existência de tendência paralela nas taxas de homicídios por cem mil habitantes entre os referidos grupos como requer uma avaliação de tratamento.

Figura 3.3 - Tendência das taxas de homicídios por cem mil habitantes dos grupos de tratamento e de controle, 2000 a 2010.



Fonte: Elaboração Própria.

Ademais, diversos programas de segurança pública foram implementados em vários estados brasileiros no período de 2000 a 2010, inviabilizando a comparação com os municípios dos demais estados, visto que resultaria em eventos heterogêneos. Caso tais municípios fossem considerados no grupo de controle, a existência de outros programas de segurança pública poderia subidentificar ou, até mesmo, sobreidentificar o impacto do INFOCRIM, causando viés à estimativa do parâmetro de interesse. Portanto, ao considerar os municípios do estado de São Paulo como área de estudo garante-se que, caso haja alterações nas leis estaduais, todos os municípios serão impactados de maneira homogênea por serem regidos pelas mesmas leis estaduais.

Além da definição dos grupos de tratamento e de controle, outra importante questão para a implementação do DID e SDID está relacionada à delimitação do período de tempo a ser avaliado. O período *ex ante* foi definido como o ano 2000, enquanto o período *ex post* refere-se ao ano de 2010. A escolha pelo referido período recai sobre o fato de que nessa década foi verificada uma queda abrupta e sistemática das taxas de criminalidade letal no estado de São Paulo, especialmente nos municípios da região metropolitana. Ademais, pelo fato de 2000 e 2010 se tratarem de anos censitários possibilita a inclusão de um número maior de variáveis de controle, bem como de municípios na amostra, o que garante algumas propriedades assintóticas do estimador mais adequado para identificar o efeito causal do INFOCRIM sobre as taxas de homicídios.

Uma vez definidos os grupos de tratamento e de controle, assim como o período de análise, a especificação econométrica para mensurar o impacto do INFOCRIM sobre a criminalidade nos municípios paulistas consiste em estimar o modelo DID por meio de um painel de dados de dois períodos. Wooldridge (2013)¹⁸ demonstra que essa especificação é equivalente à descrita na seção 3.1. Diante disso, o modelo DID é formalizado a seguir:

$$crime_{it} = \alpha_i + \beta_1 t + \beta_2 trat + \phi Infocrim_{it} + \mathbf{X}_{it}\boldsymbol{\theta} + \varepsilon_{it}, \quad (3.45)$$

Onde $t = 2000, 2010$ e i refere-se aos municípios paulistas contemplados na amostra, sendo $i = 1, \dots, 615$. α_i refere-se ao efeito não observado inerente a cada município e será tratado pelos métodos de efeitos fixos (estimador *within*¹⁹) ou efeitos aleatórios (estimador de mínimos quadrados generalizados exequíveis), levando-se em conta a escolha indicada pelo teste de Hausman (1978); t é uma *dummy* temporal em que assume o valor zero para o período *ex ante* e um para o período *ex post* ao tratamento, evidenciando a dinâmica temporal comum entre os grupos de controle e de tratamento; $trat$ é a *dummy* que identifica os grupos de tratamento e de controle, sendo que assume valor um para os municípios tratados e zero, caso contrário. Por sua vez, β_1 e β_2 são coeficientes a serem estimados que correspondem, respectivamente, ao efeito fixo do tempo e à diferença na criminalidade entre os grupos de tratamento e de controle.

Já \mathbf{X}_{it} refere-se a uma matriz contendo os diversos determinantes do crime como renda, desemprego, desigualdade, densidade demográfica - descritas na seção 3.3 - além de escolaridade, sexo, faixa etária, religião, instabilidade familiar, benefício social e outras políticas (Estatuto do desarmamento e existência de guarda municipal) apontadas pela literatura empírica como fatores que incidem sobre a dinâmica criminoso. $\boldsymbol{\theta}$ é um vetor de parâmetros associados a \mathbf{X}_{it} a serem estimados e ε_{it} corresponde ao termo de erro aleatório.

Portanto, a amostra contém 615 municípios paulistas, para os quais havia dados disponíveis, a serem analisados no intervalo compreendido entre os anos de 2000 e 2010. Os dados de homicídios e de suicídios são disponibilizados pelo Sistema de Informações de Mortalidade (SIM) da Secretaria de Vigilância à Saúde (SVS) do Ministério da Saúde

¹⁸ Maiores detalhes, consultar Wooldridge (2013), capítulo 13, seção 13.4.

¹⁹ Esse estimador é também chamado de “*within groups*” ou intergrupos.

(MS) disponíveis no site do Departamento de Informática do Sistema Único de Saúde (DATASUS), enquanto os dados socioeconômicos e demográficos foram extraídos dos Censos Demográficos de 2000 e 2010 realizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE). A variável relativa à existência de guarda municipal foi construída a partir da Pesquisa de Informações Básicas Municipais – MUNIC – 2011. Quanto às informações acerca do número de beneficiários do Programa Bolsa Família, os dados foram extraídos do Ministério do Desenvolvimento Social e Combate à Fome (MDS).

A variável dependente, taxa de homicídios por cem mil habitantes ($crime_{it}$), é definida como a razão entre o número de óbitos por homicídios (óbitos por agressões) e a população total residente em determinado município, no ano considerado, por cem mil habitantes. Os óbitos por homicídios correspondem aos códigos X85 a Y09 do capítulo XX – Causas externas de morbidade e mortalidade – da 10ª Revisão da Classificação Internacional de Doenças (CID-10), seguindo a classificação proposta pelo Ministério da Saúde²⁰. A finalidade da taxa por cem mil é permitir a comparabilidade entre locais com diferentes tamanhos de população e neutralizar o crescimento populacional, permitindo a comparação a médio e longo prazo. Formalmente, a taxa de homicídios por cem mil habitantes é calculada da seguinte forma:

$$crime_{it} = \frac{\text{Número de óbitos de residentes por agressões}_{it}}{\text{População total residente}_{it}} * 100.000 \quad (3.46)$$

O uso desta variável como *proxy* para a criminalidade violenta é justificada pelo fato de que a taxa de homicídios intencionais parece ser a melhor medida existente, uma vez que se tem uma mensuração mais acurada da ocorrência de crimes devido à alta taxa de subregistro nos demais tipos de crimes. Isso se deve ao fato de que a taxa de subregistro de óbitos por agressões é significativamente pequena por implicar perda de vida humana e, conseqüentemente, registro obrigatório no Instituto Médico Legal (JUSTUS e KASSOUF, 2008).

$Infocrim_{it}$ é uma *dummy* resultante da interação entre t e $trat$, sendo que o valor unitário corresponde aos municípios tratados após o tratamento e zero, caso contrário. O INFOCRIM, sigla do Sistema de Informações Criminais, é uma ferramenta da tecnologia da informação disponibilizado para uso em segurança pública que permite o

²⁰ A descrição de cada categoria referente aos óbitos por agressões está detalhada no Apêndice B. O grupo de causas, de acordo com a classificação do Ministério da Saúde, está disponível em: <http://tabnet.datasus.gov.br/tabdata/LivroIDB/2edrev/c09.pdf>

georreferenciamento dos crimes conforme os registros de ocorrências criminais para direcionar a atuação do policiamento especializado. Criado em 1999, o lançamento de uma rede informatizada no ano de 2000 marca o início do INFOCRIM na cidade de São Paulo. De acordo com a SSP/SP (2010), a implantação do Sistema nos 38 municípios da Região Metropolitana e nas cidades de Campinas e Santos foi concluída no ano de 2003 e, até o ano de 2010, estendida para outras 26 cidades paulistas. Sendo assim, a variável que indica o Programa ($Infocrim_{it}$), *proxy* para a política estadual de segurança pública, refere-se a uma *dummy* em que os 67 municípios que implementaram o INFOCRIM até o ano de 2010 receberam o valor unitário (1) no período *ex post* ao tratamento e os demais municípios, valor nulo (0).

Cabe notar que coeficiente dessa *dummy* (\emptyset) é o parâmetro de interesse que fornece o efeito médio do tratamento (ATE) e pode ser entendido como o impacto do INFOCRIM na criminalidade nos municípios de São Paulo. A hipótese de identificação do parâmetro \emptyset no modelo DID requer que a intervenção avaliada seja um experimento natural de tal modo que os indivíduos tratados tenham sido escolhidos aleatoriamente.

Tal hipótese é garantida por essa tese, tendo em vista que, conforme modelo teórico apresentado na seção 3.3 sob a ótica da teoria econômica do crime de Becker (1968), o Governo – representante legal da sociedade – ao aplicar rigorosamente a lei, atua sobre o aumento da probabilidade de detenção e aprisionamento (π) dos indivíduos da população N , afetando a probabilidade de fracasso do ato criminoso. Diante disso, a política de segurança pública implementada recentemente nos sessenta e sete municípios paulistas pode ser considerada um evento exógeno na oferta individual de crime, pois atua no sentido de coibir o comportamento criminoso de todos os indivíduos residentes nesses municípios.

Vale ressaltar que no caso da agregação da oferta individual de crime, a garantia de que o Programa INFOCRIM seja um quase-experimento ainda será factível, uma vez que a aplicação rigorosa da lei, via política de segurança pública INFOCRIM, consiste em uma forma de aumentar a probabilidade de detenção e aprisionamento (Π) de todos os indivíduos de χ , minimizando os prejuízos sociais da criminalidade, tendo em vista que induziria tais indivíduos a cometer um “nível ótimo de crimes”.

Além disso, a participação no tratamento será controlada por uma série de características observáveis (X_{it}) de modo que sejam suficientes para tornar a participação no Programa um evento aleatório. Como previsto pelo modelo teórico apresentado na seção 3.3, uma maior densidade demográfica pode estimular o incremento

da criminalidade. Sendo assim, é esperada que a variável densidade demográfica ($dens_{it}$), calculada como a razão da população total residente pela área territorial do espaço geográfico, no ano considerado, aumente os índices de criminalidade. Quanto à variável renda *per capita* ($renda_{it}$), definida como sendo a razão entre o somatório da renda de todos os indivíduos residentes e o número total desses indivíduos no ano considerado, a literatura econômica do crime de Becker (1968) e Ehrlich (1973) prevê um impacto ambíguo sobre a criminalidade. No que tange à taxa de desemprego (des_{it}), definida como o percentual da população de 16 anos e mais, economicamente ativa, que se encontrava desocupada, espera-se que o aumento do desemprego estimule a oferta de crimes.

No modelo econômico de Becker (1968), a educação é uma medida das habilidades do indivíduo na sua atividade produtiva (legal ou ilegal). Dessa forma, o papel da educação sobre a incidência de criminalidade também é controverso. *A priori*, espera-se que o maior nível de escolaridade reduza a criminalidade. A explicação econômica é a de que um maior nível de capital humano implica retornos esperados maiores no mercado legal e, por conseguinte, os custos de oportunidade para a prática de crimes aumentam consideravelmente. Entretanto, indivíduos qualificados podem ter maior *expertise* em cometer crimes que exijam certo grau de especialização e, logo, crimes nos quais os ganhos auferidos são expressivamente maiores (crimes de “colarinho branco”, por exemplo). Como *proxy* disponível para o nível de escolaridade, tem-se o percentual da população de 25 anos ou mais com ensino médio completo ($educ_{it}$), que é definida como a razão entre as pessoas de 25 anos ou mais que completaram o ensino médio em qualquer uma das modalidades (curso regular seriado e não seriado, Ensino de Jovens e Adultos - EJA - e supletivo) e a população total nessa faixa etária, multiplicado por 100.

A hipótese do fogo do inferno de Hirschi e Stark (1969) postula que a religião impede o comportamento criminoso, aumentando os custos da delinquência por meio da promessa de punição na outra vida. Essa hipótese foi reformulada de tal modo que o efeito causal negativo esperado entre crime e religião seria encontrado somente em áreas onde a adesão e frequência em igrejas fossem altas (hipótese da comunidade moral). Caso contrário, haveria pouco ou nenhum efeito da religiosidade sobre o comportamento criminoso dos indivíduos (STARK *et al.*, 1982; STARK, 1996). Sob a ótica da literatura econômica do crime, criminalidade e religião estariam associados devido aos custos morais que um indivíduo religioso teria em cometer um crime, afinal os membros da igreja que cometessem atos delinquentes estariam sujeitos à estigmatização e ostracismo

de sua comunidade religiosa (HULL, 2000). Portanto, o raciocínio econômico prevê que um indivíduo racional religioso seria menos propenso a cometer um ato delinquente em virtude de os custos morais associados com a realização desse ato serem maiores em comparação a uma pessoa não-religiosa, culminando em um efeito dissuasor da criminalidade (HEATON, 2006). A variável utilizada para captar o efeito da religião sobre a criminalidade é o percentual da população residente sem religião, ateu ou agnóstico ($srel_{it}$).

O perfil demográfico da população, no que diz respeito à distribuição por sexo e idade, tem sido reconhecido como um dos elementos centrais não apenas nas abordagens teóricas, mas também em inúmeras evidências empíricas. O Relatório Global sobre Homicídios da ONU, lançado em 2012, aponta que as maiores vítimas de homicídios na América Latina foram homens na faixa etária de 15 a 29 anos. Nessa vertente, o crime não é uma constante na vida do indivíduo (THORNBERRY, 1996) e está fortemente relacionado ao sexo masculino (GRAHAM e BOWLING, 1995; FLOOD-PAGE *et al.*, 2000). No contexto da escolha racional do indivíduo preconizada por Becker (1968), o jovem não possui grandes oportunidades de trabalho no mercado legal, o que estimularia a oferta de crimes por possuir baixos custos de oportunidade. Nesse contexto, a percentagem de jovens do sexo masculino ($masc_{it}$) é definida como a razão entre a população masculina na faixa etária compreendida entre 15 e 29 anos em relação à população residente total, em determinado município no ano considerado. É esperado, portanto, que um maior percentual de jovens do sexo masculino aumente os índices de criminalidade violenta.

É comumente argumentado que um maior nível de instabilidade familiar pode significar uma maior predisposição ao crime (SHAW e MCKAY, 1942). Por esse argumento, os custos morais dos potenciais criminosos são reduzidos por haver menores vínculos pessoais desencadeando, assim, uma maior atividade criminosa. Uma medida que vem sendo amplamente adotada para medir o grau de desorganização social é o percentual de lares uniparentais. Com isso, a *proxy* para a instabilidade familiar refere-se ao percentual de mães chefes de família sem fundamental e com filho menor no total de mães chefes de família ($chefe_{it}$).

Os programas sociais de transferência de renda podem ter impactos diretos sobre a criminalidade, tendo em vista que a renda adicional decorrente da participação nesses programas tende a reduzir a propensão a cometer crimes ao aumentar o custo de oportunidade da atividade ilegal dos beneficiários (BENOIT e OSBORNE, 1995;

ZHANG, 1997; IMROHOROGLU *et al.*, 2000). A variável para mensurar o impacto dos programas sociais de transferência de renda refere-se ao percentual de beneficiários do Programa Bolsa Família em relação à população total (BF_{it}).

Segundo a vertente econômica do crime de Becker (1968) e Ehrlich (1973), um maior nível de desigualdade torna os atos de violência mais atraentes em virtude de que há um maior *gap* entre o custo de oportunidade dos criminosos (renda lícita) e a renda proveniente do crime. Posto de outra forma, pelo fato de os indivíduos pobres ganharem menos ao passo que os indivíduos ricos ganham mais, o maior nível de desigualdade social pode aumentar as taxas de criminalidade. Para captar essa relação causal será utilizada a razão dos 20% mais ricos em relação aos 40% mais pobres ($r20_{it}$). Essa variável compara a renda *per capita* média dos indivíduos pertencentes ao quinto dos indivíduos mais ricos nessa distribuição com a renda *per capita* média dos indivíduos pertencentes aos dois quintos mais pobres.

Com relação às políticas de desarmamento da população, tampouco há consenso na literatura dos seus efeitos sobre a criminalidade. Por um lado, os defensores da hipótese de “mais armas, mais crimes” entendem que a proibição e restrição da posse e porte de armas é uma medida importante contra a atividade criminosa (DUGAN, 2001, por exemplo). No outro extremo, os defensores da hipótese de “mais armas, menos crimes” entendem que ao retirar armas de circulação, haveria redução do custo esperado da atividade ilegal (LOTT JR. e MUSTARD, 1997, por exemplo). A *proxy* para difusão de armas de fogo nos municípios consistirá na proporção de suicídios por perfuração de arma de fogo - Lesões autoprovocadas intencionalmente por arma de fogo, categorias X72 a X74 da CID 10 - em relação ao total de suicídios - Lesões autoprovocadas intencionalmente, categorias X60 a X84 da CID 10 - (*saf*)²¹. Essa *proxy* foi utilizada por Kleck (1997), Moody e Marvell (2002), Cerqueira (2010) e outros e validada por Kleck (2004) ao investigar 25 diferentes *proxies* utilizadas na literatura para a prevalência de armas de fogo nas localidades, chegando à conclusão de que a única *proxy* válida seria exatamente a *saf*. Por intermédio do uso da variável *saf* pretende-se captar o possível efeito da política de *buy back* realizada em âmbito nacional, o Estatuto do Desarmamento.

As guardas municipais podem ser inseridas no contexto do policiamento comunitário (SKOLNICK e BAYLEY, 2002) e do policiamento orientado para a solução

²¹ A descrição de cada categoria referente às lesões autoprovocadas intencionalmente está detalhada no Apêndice C.

de problemas (GOLDENSTEIN, 1979) no sentido de que atuação da guarda municipal pode ter efeito indireto sobre a prevenção da violência e da criminalidade letal, uma vez que sua atribuição é a de proteção do patrimônio público municipal. A existência de guarda municipal com porte de arma de fogo ($garma_{it}$) pode inibir indiretamente a criminalidade letal, tendo em vista que, ao exercer sua função, retiram o excesso de demanda das polícias estaduais para que estas exerçam seu papel de combate ao crime. A variável ($garma_{it}$) é uma *dummy* que assume valor unitário para os municípios paulistas com existência de guarda municipal com porte de arma de fogo e zero para os demais municípios.

O Quadro 3.3 resume as informações sobre as variáveis descritas e as respectivas fontes estabelecendo a conexão entre os modelos teórico e empírico por meio das variáveis teóricas e seus respectivos sinais esperados no âmbito do modelo empírico.

Quadro 3.3 - Resumo das variáveis utilizadas no modelo empírico.

	Variável Teórica	Notação	Proxy	Unidade	Sinal Esperado	Arcabouço teórico/empírico	Fonte
Variável Dependente	Crime	Crime	Taxa de homicídio por cem mil habitantes	Taxa por cem mil	_____	Becker (1968); Ehrlich (1973)	DATASUS/MS
Variáveis Explicativas	<i>Deterrence effects</i> (probabilidade de detenção e aprisionamento)	Infocrim	Programa INFOCRIM	0 ou 1	Negativo	Becker (1968); Ehrlich (1973)	Variável Construída
	Retorno esperado / Custo de oportunidade	Dens	Densidade populacional	População por km ²	Positivo	Glaeser e Sacerdote (1999); Kelly (2000)	Censo/IBGE
	Retorno esperado / Custo de Oportunidade	Renda	Renda <i>per capita</i>	R\$	Positivo/ Negativo	Becker (1968); Ehrlich (1973)	Censo/IBGE
	Custo de oportunidade	Des	Taxa de desemprego	Taxa	Positivo	Becker (1968); Ehrlich (1973); Merton (1938)	Censo/IBGE
	Custo de oportunidade	Educ	População com ensino médio completo	Proporção	Positivo/ Negativo	Becker (1968)	Censo/IBGE
	Custo moral	Srel	População sem religião	Proporção	Positivo	Hull (2000); Heaton (2006)	Censo/IBGE
	Custo de oportunidade	Masc	Jovens do sexo masculino	Proporção	Positivo	Graham e Bowling (1995); Thornberry (1996); Flood-Page <i>et al.</i> (2000)	Censo/IBGE
	Desorganização social	Chefe	Mulheres chefes de família	Proporção	Positivo	Shaw e McKay (1942)	Censo/IBGE
	Custo de oportunidade	BF	Beneficiários do Bolsa Família	Proporção	Negativo	Benoit e Osborne (1995); Zhang (1997); Imrohoroglu <i>et al.</i> (2000)	MDS
	Desigualdade social	r20	Razão 20% mais ricos / 40% mais pobres	Proporção	Positivo	Ehrlich (1973); Kelly (2000)	Censo/IBGE
	<i>Deterrence effects</i>	Saf	Suicídios por armas de fogo	Proporção	Positivo/ Negativo	Lott Jr. e Mustard (1997); Dugan (2001)	DATASUS/MS
<i>Deterrence effects</i>	Garma	Existência de guarda municipal com porte de arma de fogo	0 ou 1	Negativo	Skolnick e Bayley (2002)	MUNIC/IBGE	

Fonte: Elaboração Própria.

A Tabela 3.1 apresenta a estatística descritiva dos dados que serão utilizados na especificação econométrica para os grupos de tratamento e de controle da taxa de homicídios por cem mil habitantes, bem como dos demais determinantes da criminalidade para o período *ex ante* à implementação efetiva do INFOCRIM. Tais estatísticas são úteis para verificar que, uma vez controladas para os determinantes da criminalidade, a participação no tratamento não dependeria dos resultados, garantindo a hipótese de independência condicional.

Já na Tabela 3.2 são apresentadas as correlações simples entre as variáveis incluídas no modelo. Pode-se observar que não existe multicolinearidade entre as variáveis, isentando o modelo do referido problema econométrico. Sendo assim, é possível calcular o efeito marginal de cada um dos determinantes da criminalidade, bem como o efeito médio do tratamento.

Tabela 3.1 - Estatística descritiva dos dados no período *ex ante* ao tratamento, 2000.

Grupo de tratamento					
Variáveis	Observações	Média	Desvio-padrão	Mínimo	Máximo
Crime	67	38.95	25.01	0.00	108.66
Infocrim	67	0.00	0.00	0.00	0.00
Dens	67	1912.74	2819.39	33.72	11630.80
Renda	67	758.16	267.26	368.46	1639.93
Des	67	18.41	4.03	6.89	26.00
Educ	67	10.21	0.48	9.02	11.47
Srel	67	0.08	0.04	0.02	0.24
Masc	67	0.14	0.01	0.12	0.17
Chefe	67	7.30	1.08	5.09	9.91
Bf	67	0.00	0.00	0.00	0.00
r20	67	10.86	3.69	6.43	32.53
Saf	67	0.17	0.21	0.00	1.00
Garma	67	0.55	0.50	0.00	1.00
Grupo de controle					
Crime	548	10.45	14.12	0.00	92.99
Infocrim	548	0.00	0.00	0.00	0.00
Dens	548	78.24	192.84	3.93	2682.08
Renda	548	527.42	147.24	203.69	1235.29
Des	548	11.91	4.34	1.69	29.72
Educ	548	10.13	0.72	6.91	12.59
Srel	548	0.04	0.03	0.00	0.20
Masc	548	0.14	0.01	0.09	0.23
Chefe	548	5.47	1.53	2.07	11.69
Bf	548	0.00	0.00	0.00	0.00
r20	548	9.62	2.69	4.58	25.74
Saf	548	0.07	0.22	0.00	1.00
Garma	548	0.09	0.29	0.00	1.00

Fonte: Elaboração Própria.

Tabela 3.2 - Matriz de correlação das variáveis do modelo empírico.

	crime	infocrim	Dens	Renda	Des	Educ	srel	Masc	chefe	bf	r20	saf	garma
Crime	1.00												
Infocrim	0.08	1.00											
Dens	0.32	0.38	1.00										
Renda	0.04	0.33	0.24	1.00									
Des	0.34	-0.08	0.19	-0.21	1.00								
Educ	-0.09	0.08	-0.01	0.15	-0.14	1.00							
Srel	0.31	0.23	0.24	0.08	0.21	-0.06	1.00						
Masc	0.10	-0.05	0.04	-0.11	0.23	-0.16	0.23	1.00					
Chefe	0.10	0.34	0.18	0.25	-0.25	0.07	0.32	-0.07	1.00				
Bf	-0.04	0.13	-0.04	0.11	-0.53	0.22	0.20	-0.15	0.56	1.00			
r20	0.14	0.03	0.03	0.11	0.45	-0.20	0.10	0.00	-0.16	-0.39	1.00		
Saf	0.06	0.06	0.08	0.12	0.13	-0.04	0.04	0.00	-0.03	-0.09	0.10	1.00	
Garma	0.23	0.32	0.36	0.42	0.15	-0.07	0.16	0.07	0.15	-0.10	0.09	0.11	1.00

Fonte: Elaboração Própria

Além dos diversos determinantes da dinâmica criminosa a serem considerados, Sickles e Williams (2006) argumentam que recentes pesquisas teóricas e empíricas sugerem que as interações são particularmente relevantes para a decisão de participação no crime, se o estigma associado à prisão age como um efeito dissuasor. Diante dessa ressalva, deve-se levar em consideração a possibilidade de interação espacial entre os municípios vizinhos ou migração da criminalidade e, nesse caso, a hipótese SUTVA não será garantida. A estratégia para testar a hipótese SUTVA consiste em verificar os resíduos da equação (3.45) para a existência (ou não) da dependência espacial, utilizando, por exemplo, o teste de autocorrelação espacial I de Moran.

Uma vez comprovada a existência de autocorrelação espacial nos resíduos da regressão DID, a hipótese SUTVA deve ser flexibilizada de maneira a modelar a interação espacial entre os municípios paulistas e, assim, identificar o efeito causal do INFOCRIM (\emptyset). Para tanto, o método SDID é capaz de modelar o transbordamento da criminalidade. Assim, considerando o grupo de controle como um bom contrafactual e a flexibilização da premissa SUTVA na presença de dependência espacial, uma especificação geral do SDID capaz de modelar todas as formas existentes de autocorrelação espacial pode ser feita por meio de uma extensão da equação (3.45) do DID, como segue:

$$crime_{it} = \alpha_i + \beta_1 t + \beta_2 trat + \emptyset Infocrim_{it} + \rho W crime_{it} + \delta W Infocrim_{it} + X_{it} \theta + W X_{it} \tau + \xi_{it} \quad (3.47a)$$

$$\xi_{it} = \lambda W \xi_{it} + \varepsilon_{it} \quad (3.47b)$$

Em que a variável $W crime_{it}$ refere-se à criminalidade defasada espacialmente e o parâmetro subjacente ρ pode ser entendido como o transbordamento da criminalidade; $W Infocrim_{it}$ consiste na defasagem espacial do Programa INFOCRIM e δ é o coeficiente a ser estimado e pode ser entendido como o efeito indireto do Programa para todos os municípios, tratados e não tratados, condicionado a um vizinho tratado; $W X_{it}$ é uma matriz dos determinantes da criminalidade defasados espacialmente e o vetor de coeficientes τ mensura os transbordamentos espaciais desses determinantes; $W \xi_{it}$ referem-se aos transbordamentos não modelados, sendo λ o parâmetro de erro autorregressivo espacial. W é a matriz de ponderação espacial (615x615) que melhor define o arranjo espacial entre os 615 municípios paulistas, sendo W , por pressuposição, constante ao longo do tempo. O restante da notação permanece a mesma. Novamente, o

interesse recai sobre o parâmetro \emptyset que identifica o impacto do INFOCRIM sobre a criminalidade nos municípios de São Paulo.

A equação (3.47) é conhecida na literatura econométrico-espacial como modelo de Manski. A partir das diferentes restrições sobre os parâmetros espaciais da equação (3.47), a taxonomia dos diversos modelos espaciais pode ser assim definida: i) se $\rho \neq 0$, $\tau = 0$ e $\delta = \lambda = 0$, tem-se o modelo SDID-SAR; ii) se $\rho = \delta = 0$, $\tau = 0$ e $\lambda \neq 0$, trata-se de um modelo SDID-SEM; iii) se $\rho \neq 0$, $\tau = 0$, $\delta = 0$ e $\lambda \neq 0$, o modelo a ser estimado é um SDID-SAC; iv) se $\rho = 0$, $\tau \neq 0$, $\delta \neq 0$ e $\lambda = 0$, o modelo é um SDID-SLX; v) se $\rho \neq 0$, $\tau \neq 0$, $\delta \neq 0$ e $\lambda = 0$, tem-se o modelo SDID-SDM; vi) se $\rho = 0$, $\tau \neq 0$, $\delta \neq 0$ e $\lambda \neq 0$, refere-se a um modelo SDID-SDEM e, por fim, vale destacar que vii) se $\rho = \delta = \lambda = 0$ e $\tau = 0$ tem-se o modelo DID convencional (equação 4.1).

A taxonomia dos modelos SDID será estimada pelo estimador inter-grupos adaptado para resolver a provável endogeneidade decorrente da simultaneidade das variáveis *gamma* e *Wcrime*, proposto por Elhorst (2003). O estimador proposto por Elhorst (2003) estima a equação diminuída pela média (*demeaned equation*) por máxima verossimilhança usando os resíduos da equação diferenciada, em vez de estimar por MQO como no estimador inter-grupos convencional. O procedimento de Elhorst (2003) permite resolver a simultaneidade, inclusive aquela derivada da defasagem espacial (*Wcrime*).²² Cabe ressaltar que os modelos DID e SDID serão estimados com erros-padrão robustos (matriz de variância-covariância de White) para tratar a possível heterocedasticidade existente.

Após a estimação dos diversos modelos SDID, o modelo que melhor identifica o parâmetro de interesse \emptyset é aquele que apresentar melhor ajuste definido pelo critério de informação bayesiano (BIC), pois o referido critério possui melhor desempenho para grandes amostras e, ao penalizar mais os parâmetros adicionais incluídos no modelo em relação ao critério de informação de Akaike (AIC), tende a selecionar modelos mais parcimoniosos (ENDERS, 2010).

3.5 A CRIMINALIDADE EVITADA PELO INFOCRIM

²² Para maiores detalhes técnicos desse estimador, ver Elhorst (2003).

Muitos analistas têm buscado avaliar o impacto de políticas de segurança pública de modo a subsidiar o planejamento e a gestão dos programas implementados, mas pouca importância tem sido dada à mensuração de vidas poupadas em decorrência da implementação dos programas de segurança pública. Diante disso, caso seja verificado que o INFOCRIM tenha contribuído para a redução da taxa de homicídios nos municípios do estado de São Paulo, torna-se possível mensurar a quantidade de homicídios evitados pelo INFOCRIM a partir da metodologia proposta a seguir.

Considerando que o ATE (efeito médio do tratamento) foi identificado garantindo todas as premissas necessárias para estimar \emptyset de forma consistente e eficiente, é pertinente afirmar que \emptyset seja o impacto do INFOCRIM sobre a criminalidade. Caso \emptyset seja negativo, esse coeficiente pode ser interpretado como a taxa homicídios por cem mil habitantes evitada pelo INFOCRIM, uma vez que a variável dependente analisada é a taxa de homicídios por cem mil habitantes nos municípios paulistas. Formalmente, o número de crimes violentos evitado pelo INFOCRIM pode, então, ser calculado da seguinte forma:

$$\text{Número de Crimes Evitados}_{\text{INFOCRIM}} = \frac{\emptyset * \text{Popmunic}}{100.000} \quad (3.48)$$

Onde: \emptyset é o ATE ou o impacto do Programa sobre a taxa de homicídios; *Popmunic* é a população total dos 615 municípios contemplados na amostra.

A mensuração do número de vidas poupadas em decorrência da implementação do Programa torna-se importante para subsidiar as decisões dos *policy makers* quanto à continuidade do Programa ou à busca de políticas alternativas com vistas a gerar maior bem-estar à sociedade.

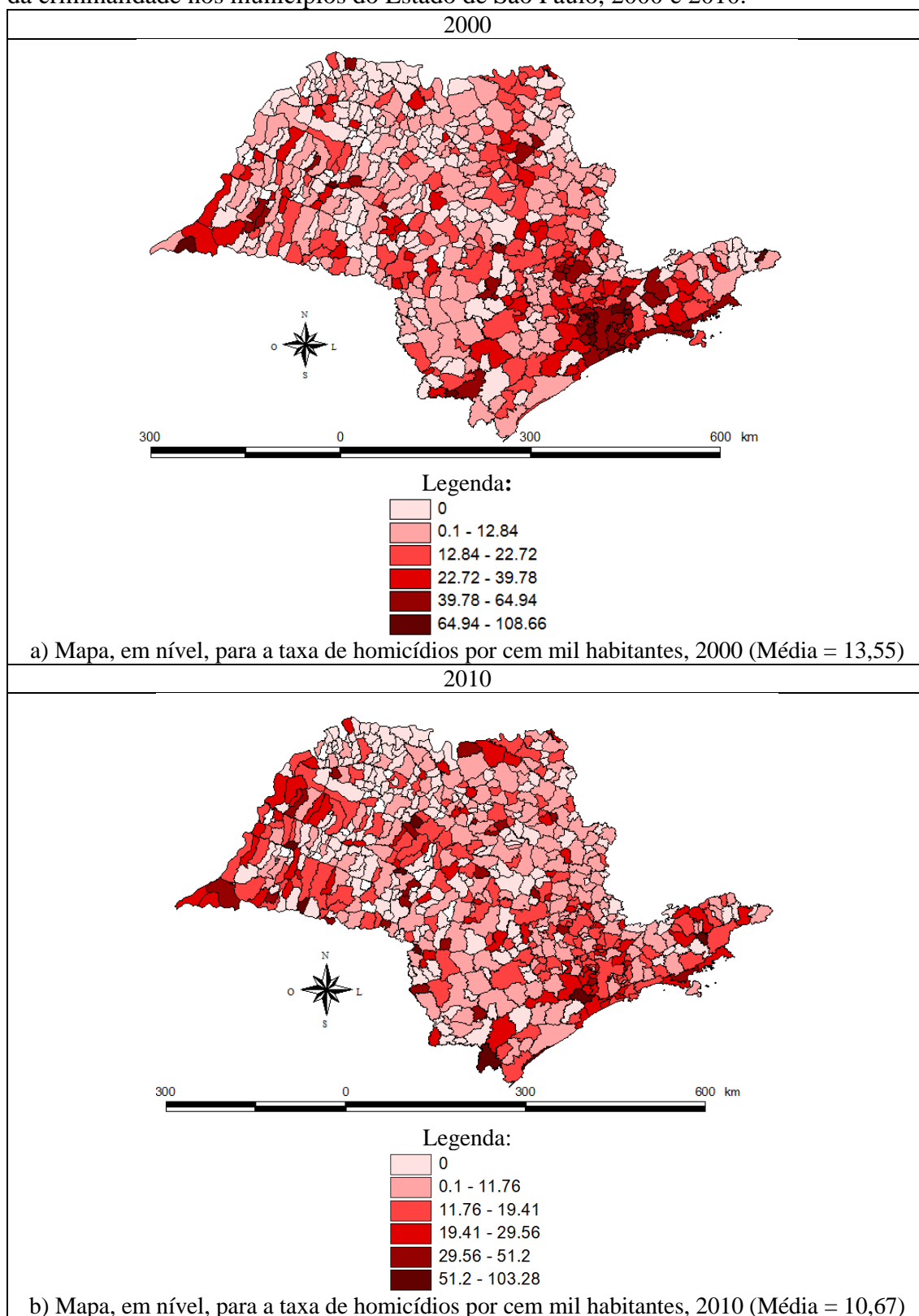
3.6 ANÁLISE EXPLORATÓRIA DOS DADOS ESPACIAIS

Sob a forma de gráficos e mapas, a análise exploratória de dados espaciais (AEDE) permite detectar padrões de associação espacial e *clusters* por intermédio da medida de autocorrelação I de Moran (global e local), de modo a sugerir se as interações espaciais são importantes para explicar o fenômeno da criminalidade em São Paulo para os anos de 2000 e 2010. Conforme o procedimento de Baumont (2004), a matriz de ponderação espacial escolhida foi a de k -vizinhos mais próximos, sendo $k=5$. Os resultados indicaram que, para os anos analisados, os maiores valores de I de Moran foram obtidos com a referida matriz.

Com vistas a compreender a dinâmica espacial da criminalidade nos municípios paulistas, a Figura 3.4 (a e b) evidencia a queda da criminalidade. A distribuição espacial da taxa de homicídios nos anos 2000 e 2010 (Figuras 3.4a e 3.4b, respectivamente) mostra que a região da Grande São Paulo era composta por municípios com expressivos índices de criminalidade no ano de 2000, enquanto no ano de 2010 pode ser observada uma reversão da dinâmica das ocorrências de homicídios na Região Metropolitana de São Paulo.

Entretanto, vale observar que, em 2010, municípios como Lagoinha na Região Administrativa (RA) de São José dos Campos; Barra do Turvo e Ilha Comprida na RA de Registro; Bocaina, Ubirajara e Sabino na RA de Bauru; Sarutaiá e Riversul na RA de Itapeva - Regiões Administrativas vizinhas às Regiões onde o INFOCRIM foi implementado - aumentaram seus níveis de criminalidade no decorrer da década de 2000. É perceptível também o aumento da criminalidade no decorrer da década na região oeste do estado, mais especificamente, em Guaraçai na RA de Araçatuba; Nantes, Flora Rica e Estrela do Norte na RA de Presidente Prudente; Santa Cruz da Esperança na RA de Ribeirão Preto, dentre outros, sugestionando possível reestruturação da atividade criminosa no estado de São Paulo. Uma observação deve ser feita à criminalidade na cidade de Juquitiba na RMSP que, apesar de ser um dos municípios contemplados com o INFOCRIM, aumentou sua criminalidade no período em questão.

Figura 3.4 – Mapas para a taxa de homicídios por cem mil habitantes, em nível (a e b), da criminalidade nos municípios do Estado de São Paulo, 2000 e 2010.



Fonte: Elaboração Própria.

Dando continuidade à AEDE, a estatística *I* de Moran univariada (segunda e terceira colunas da Tabela 3.3 em 2000 e 2010) - que relaciona uma variável específica à sua defasagem espacial - revela a existência de um padrão de similaridade na taxa de homicídio por cem mil habitantes, logo, um padrão espacial de concentração. Ou seja, municípios com alta criminalidade são vizinhos de municípios com altos níveis de crime e municípios onde a criminalidade é baixa são circunvizinhos de municípios com criminalidade reduzida em ambos os anos. Esse mesmo padrão de concentração espacial é evidenciado para os demais determinantes da criminalidade no contexto univariado: municípios que possuem altos (baixos) níveis de renda, desigualdade, desemprego, desorganização familiar, população masculina jovem, escolaridade, densidade demográfica, população sem religião, proporção da população assistida com o Bolsa Família são vizinhos de municípios com altos (baixos) níveis das respectivas defasagens espaciais. Quanto à *proxy* para o ED, apesar de manter o padrão de concentração espacial, o mesmo não se mostrou significativo.

A estatística *I* de Moran pode ser aplicada também a uma análise bivariada com vistas à compreensão do comportamento da taxa de homicídios por cem mil habitantes em relação aos demais determinantes da criminalidade (quarta e quinta colunas da Tabela 3.3 para 2000 e 2010). Os resultados indicam um padrão de concentração espacial entre a taxa de homicídios por cem mil habitantes com os demais determinantes da criminalidade, evidenciando que um município com alta taxa de crime é circundado por municípios com altos níveis de renda, desigualdade, desemprego, desorganização familiar, população masculina jovem, escolaridade, densidade demográfica, população sem religião e proporção da população assistida com o Bolsa Família, enquanto que municípios com baixa taxa de crime encontravam-se rodeados por municípios com reduzidos níveis dos citados determinantes da criminalidade. Vale observar que a maioria dos valores de *I* de Moran uni e bivariados foram positivos e estatisticamente significativos. O padrão negativo é observado apenas para a taxa de crimes e ED, embora não significativo.

Tabela 3.3 – Estatísticas *I* de Moran uni e bivariada, 2000 e 2010.

Variáveis	<i>I</i> de Moran univariado		<i>I</i> de Moran bivariado	
	2000	2010	2000	2010
Crime	0,473***	0,111***	—	—
Dens	0,704***	0,709***	0,479***	0,120***
Renda	0,312***	0,296***	0,189***	0,024*
Des	0,486***	0,397***	0,037***	0,152***
Educ	0,168***	0,278***	0,160***	0,117***
Srel	0,532***	0,418***	0,336***	0,138***
Masc	0,214***	0,084***	0,177***	0,035**
Chefe	0,055***	0,058***	0,300***	0,120***
BF	—	0,490***	—	0,049***
r20	0,275***	0,227***	0,116***	0,105***
Saf	0,002	0,026	-0,006	-0,008

Fonte: Elaboração Própria.

Notas: * *p*-valor<0,10; ** *p*-valor<0,05; *** *p*-valor<0,01; não significativo *p*-valor>0,10.

A estatística *I* de Moran não foi calculada para a variável *BF* no ano 2000 em virtude de a mesma ser constituída por valores nulos.

Para o cômputo da estatística de autocorrelação espacial para as variáveis binárias INFOCRIM e existência de guarda municipal com porte de arma de fogo (*garma*) deve-se utilizar o teste *join count*. Essa estatística é útil para verificar a existência de *clusters* ou dispersão em dados binários sob a hipótese nula de aleatoriedade espacial. Caso seja comprovada a existência de dependência espacial, a estatística de *joint count*, assim como o teste *I* de Moran, revela o padrão de autocorrelação espacial de dispersão ou de concentração.

Quanto ao INFOCRIM e à existência de guarda municipal com porte de arma de fogo, é importante verificar se tais políticas de segurança pública estadual e municipal, respectivamente, seguem um padrão espacialmente dependente ou se, por outro lado, refere-se a um tipo de política distribuída aleatoriamente entre os municípios paulistas. Por intermédio da análise dos resultados constantes na Tabela 3.4, em ambos os casos, pode-se perceber a existência de um padrão de concentração evidenciado pelo expressivo número de junções do tipo *Black-Black* (BB)²³, estatisticamente significativas. Em outras palavras, existe um padrão de agrupamento na implementação tanto do INFOCRIM quanto da guarda municipal com porte de arma de fogo nos municípios paulistas, fazendo com que essas variáveis sejam espacialmente autocorrelacionadas.

²³ Chama-se de região *black* (B), a região que possui a característica analisada e *white* (W), a região sem a característica. No presente caso, *black* diz respeito aos municípios onde o INFOCRIM foi implementado e os municípios com existência de guarda municipal com porte de arma de fogo.

Tabela 3.4 – Teste *join count* para a autocorrelação espacial.

Variáveis	Tipo de Junção	Número de Junções	Média	Desvio-padrão	<i>p-valor</i>
Infocrim 2010	BB	97	18,005	3,777	0,000
Garma 2000		124	29,764	4,688	0,000
Garma2010		141	34,088	4,96	0,000
Infocrim 2010	BW	142	298,990	7,555	0,977
Garma 2000		183	370,472	9,376	0,987
Garma 2010		178	391,824	9,921	0,964

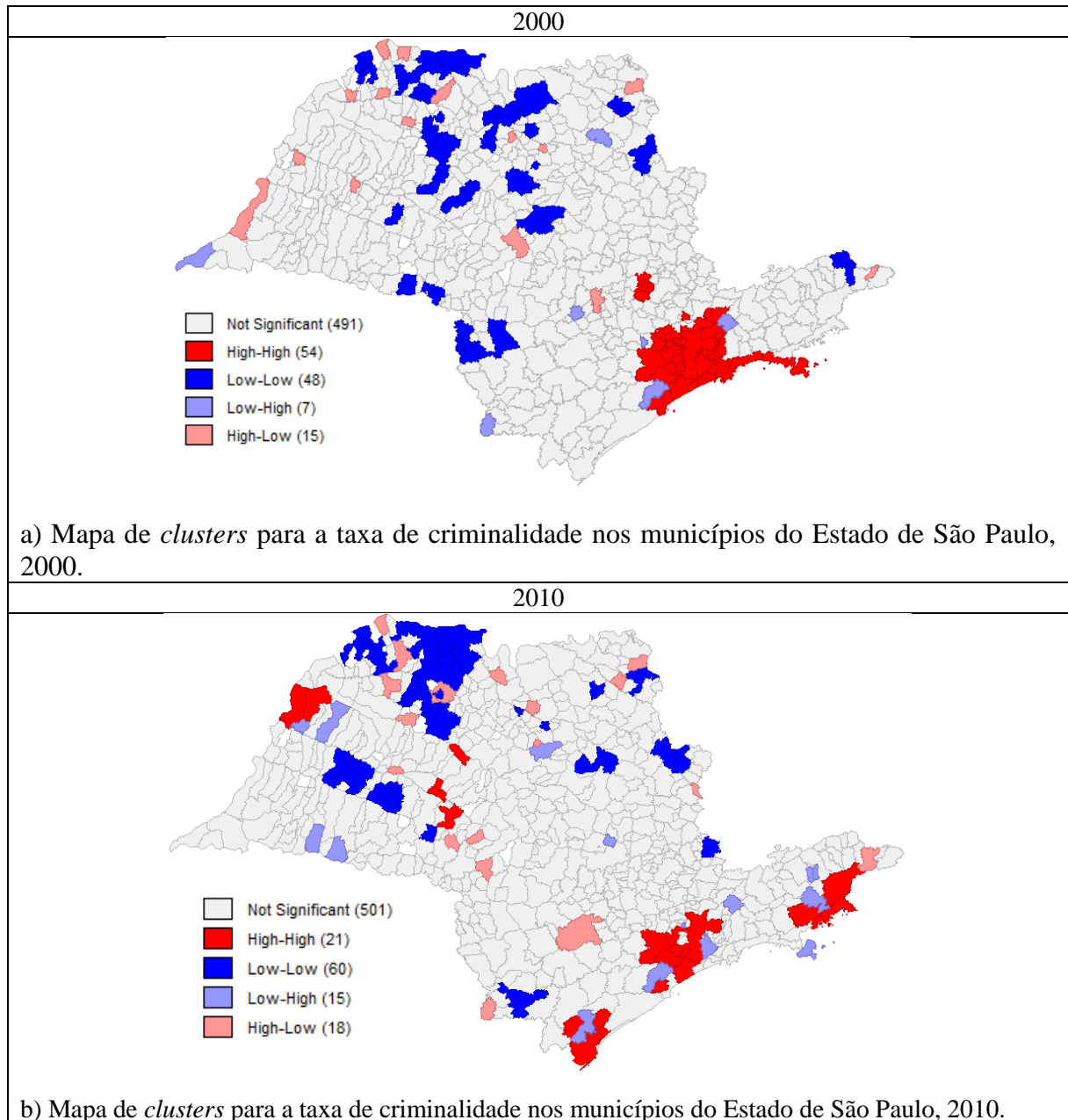
Fonte: Elaboração Própria.

Nota: A estatística *join count* para a variável *Infocrim* em 2000 pode ser calculada, pois é uma variável com todos os valores iguais a zero no referido ano.

Apesar de úteis, as estatísticas de autocorrelação global não têm a capacidade de identificar a existência de *clusters*, conforme destacado por Almeida (2012). Entretanto, Justus e Kassouf (2008) alertam para o fato da possibilidade de existência de aglomerações da atividade criminosa, ao se trabalhar com dados agregados em nível regional. Diante do exposto, será adotada a estatística *I* de Moran local, um dos Indicadores Locais de Associação Espacial (*LISA*), apresentada na forma de mapas de *clusters*, para verificar a existência (ou não) de um padrão de distribuição espacial da taxa de homicídios por cem mil habitantes nos municípios paulistas. A estatística *I* de Moran local pode ser interpretada da seguinte maneira: valores positivos de I_{it} significam que há *clusters* espaciais com valores similares (alto ou baixo); valores negativos significam que há *clusters* espaciais com valores diferentes entre os municípios e seus vizinhos.

Por intermédio da visualização da Figura 3.5 (a e b), é possível notar que, de fato, existem diferenças na distribuição espacial dos *clusters* presentes na taxa de criminalidade no período compreendido entre os anos de 2000 e 2010. Conforme pode ser observado, os municípios da RMSP, da RA de Santos e os municípios de Americana, Hortolândia, Monte-Mor, Nova Odessa e Sumaré na RA de Campinas faziam parte do *cluster* AA em 2000 (Figura 3.5a), ou seja, esses municípios possuíam altos índices de criminalidade e eram circunvizinhos de municípios com elevados níveis de taxas de homicídios por cem mil habitantes. Já em 2010 (Figura 3.5b), é perceptível a redução do número de municípios pertencentes ao *cluster* AA: de 54 municípios em 2000 para 21 em 2010. Há de se atentar que o número de municípios da RMSP presentes no *cluster* AA reduziu consideravelmente em 2010, o que pode ser decorrente da implementação do INFOCRIM no início da década.

Figura 3.5 - Mapas de *clusters* e de significância para a taxa de criminalidade por cem mil habitantes nos municípios do Estado de São Paulo, 2000 e 2010.



Fonte: Elaboração Própria.

De modo oposto, os municípios Cajati, Cananeia e Pariquera-Açu na RA de Registro; Cunha, Natividade da Serra e Ubatuba na RA de São José dos Campos e Andradina e Castilho na RA de Araçatuba que, em 2000, não se apresentavam como pertencentes a nenhum tipo de aglomeração espacial passam, em 2010, a fazer parte do *cluster* AA. Esse resultado sugere uma possível migração da criminalidade para essas RA do estado de São Paulo a partir da implantação do INFOCRIM.

Com relação ao *cluster* BB, observa-se que esse padrão de aglomeração espacial se mantém na região Norte do estado, assim como na RA Central no decorrer do período analisado.

Cabe observar que a RA de São José do Rio Preto situada na região Norte do estado passa a ter um número maior de municípios pertencentes a esse tipo de *cluster*, indicando que os municípios da referida RA reduziram, relativamente aos demais municípios do estado, seus níveis de criminalidade como um todo.

Diante dos resultados das estatísticas espaciais, há evidências de que a criminalidade no estado de São Paulo exibe um padrão de concentração espacial, sugerindo que as interações espaciais são importantes para explicar o fenômeno da criminalidade nos municípios paulistas no período considerado. Resultados similares de evidências de um padrão espacial de aglomeração da criminalidade também foram encontradas por Sartoris Neto (2000) para os distritos policiais da cidade de São Paulo, Peixoto (2003) para o município de Belo Horizonte/MG, Almeida *et al.* (2005) para 754 municípios de Minas Gerais e Justus e Santos Filho (2011) para as microrregiões brasileiras.

4 ANÁLISE DO IMPACTO DO INFOCRIM SOBRE A CRIMINALIDADE

A criminalidade no estado de São Paulo reduziu consideravelmente na última década. Alguns pesquisadores (KAHN, 2007; GOERTZEL e KAHN, 2009; NOBREGA JR *et al.*, 2009) e o Governo do estado de São Paulo creditam a redução da criminalidade à política de segurança implementada pelo Estado. Por outro lado, Mello e Schneider (2010) atribuem a queda na criminalidade à melhoria dos indicadores socioeconômicos e demográficos na última década. Em decorrência da ausência de consenso, a hipótese testada consiste em inferir se o INFOCRIM foi preponderante para a reversão dos índices de criminalidade no estado de São Paulo por intermédio de uma análise econométrica de avaliação de tratamento.

Para empreender tal análise, o grupo de tratamento contempla os 67 municípios paulistas onde foi implementado o INFOCRIM, enquanto os demais municípios do estado de São Paulo constantes na amostra (548) fazem parte do grupo de controle. Portanto, considerando uma amostra de 615 municípios paulistas, o impacto do INFOCRIM será mensurado por meio dos modelos de DID e SDID, estimados para um painel de dados de dois períodos na década de 2000 (período *ex ante*) a 2010 (período *ex post*). Há de se atentar que diversos determinantes da criminalidade apontados pela teoria econômica do crime também serão utilizados como controle na análise.

Com o intuito de verificar qual estimador modela de forma mais adequada os efeitos não observados inerentes a cada município (α_i), o modelo DID foi estimado, considerando os efeitos aleatórios por meio dos Mínimos Quadrados Generalizados Exequíveis (MQGE) e os efeitos fixos por meio do estimador *within*. Tais efeitos não observados estão relacionados a características específicas de cada município que influenciam a criminalidade letal, porém não são possíveis de serem mensuradas. Essas características podem ser decorrentes da diversidade da malha municipal do estado de São Paulo no que tange à extensão geográfica e número de municípios existentes, o que acarreta heterogeneidade cultural, de crenças, valores e institucional aos municípios paulistas.

Os resultados do modelo DID estimado por efeitos aleatórios e por efeitos fixos encontram-se reportados na Tabela 4.1. Esses modelos foram comparados por meio do teste de Hausman (1978) com o intuito de verificar qual dos dois modelos trata de forma adequada os efeitos não observados, garantindo eficiência e consistência às estimativas (parte inferior da Tabela 4.1). A hipótese nula do teste de que as estimativas por efeitos aleatórios seriam consistentes é rejeitada a 1% de significância. Assim, o modelo estimado por efeitos fixos é

atestado como o mais adequado e, portanto, o estimador *within* será empregado para estimar o impacto do INFOCRIM sobre a criminalidade.

Tabela 4.1 – Resultado das estimações dos modelos de diferenças em diferenças, DID.

Variáveis	Efeitos Aleatórios	Efeitos Fixos	Teste de sensibilidade
Tempo (t)	-2,401 (2,547)	7,221 (4,495)	8,142* (4,509)
Tratamento (trat)	13,130*** (2,325)	—	—
Infocrim	-10,306*** (2,218)	-7,986*** (2,837)	—
Densidade (Dens)	0,002*** (0,000)	-0,033*** (0,006)	-0,039*** (0,005)
Renda (Renda)	0,004 (0,004)	0,014 (0,010)	0,016 (0,010)
Desemprego (Des)	0,765*** (0,128)	0,431** (0,209)	0,573*** (0,204)
Educação (Educ)	-0,077 (0,097)	-0,604*** (0,222)	-0,710*** (0,220)
Sem religião (Srel)	67,302*** (13,262)	24,966 (26,276)	28,213 (26,401)
Homens Jovens (Masc)	-9,705 (24,053)	-13,619 (36,922)	-11,524 (37,125)
Lar uniparental (Chefe)	0,246* (0,134)	0,020 (0,185)	-0,009 (0,186)
Bolsa Família (BF)	79,086** (34,834)	-51,574 (41,344)	-44,387 (41,501)
Desigualdade (r20)	-0,056 (0,188)	-0,736* (0,412)	-0,793* (0,414)
Desarmamento (Saf)	0,081 (1,991)	-1,561 (2,673)	-1,519 (2,688)
Guarda Municipal (Garma)	4,406*** (1,361)	-18,923** (7,417)	-21,295*** (7,411)
Constante	-1,653 (3,992)	30,997*** (8,504)	32,246*** (8,541)
Teste de Hausman	58,70***		
R ²	0,115	0,223	0,213
BIC	—	8916,734	8925,7
Autocorrelação espacial nos resíduos	—	sim	—
Número de observações	1230	1230	1230

Fonte: Elaboração Própria.

Notas: Erros-padrão robustos entre parênteses.

*** p-valor<0,01, ** p-valor<0,05, * p-valor<0,10

Justus e Kassouf (2007) também encontraram evidências de que o modelo de efeitos fixos seja mais adequado que o modelo de efeitos aleatórios na análise realizada da influência do mercado de drogas ilícitas sobre a criminalidade nos estados brasileiros. Segundo os autores, no caso da criminalidade, os efeitos específicos não observados são potencialmente correlacionados com as variáveis de controle e, portanto, a estimação por efeitos fixos parece ser a mais adequada.

Por meio dos resultados obtidos com o modelo estimado por efeitos fixos (terceira coluna da Tabela 4.1), controlando por diversos determinantes da criminalidade apontados pela teoria econômica do crime, é possível inferir que há indícios de que a implementação do Programa foi exitosa em reduzir a criminalidade violenta no estado de São Paulo. O coeficiente do INFOCRIM, negativo e estatisticamente significativo, indica que o Programa pode ser considerado um fator redutor das estatísticas letais nos municípios paulistas no período analisado. O efeito médio do tratamento aponta uma redução de 7,986 na taxa de homicídios por cem mil habitantes nos municípios paulistas entre os anos de 2000 e 2010. Cabe observar que, caso fosse considerado o estimador de efeitos aleatórios para controlar os efeitos não observados inerentes a cada município, o efeito médio do INFOCRIM estaria sobreidentificado, levando à inferência errônea de que o INFOCRIM seria responsável por uma redução de 10,306 sobre as taxas de homicídios por cem mil habitantes.

Com o intuito de atestar a importância do INFOCRIM sobre a trajetória da criminalidade no estado de São Paulo, um teste de sensibilidade foi realizado (última coluna da Tabela 4.1). Tal teste consistiu em omitir a variável INFOCRIM da regressão estimada por efeitos fixos com vistas a verificar se, independentemente do INFOCRIM, a criminalidade no estado reduziria ao longo do tempo. A política é importante para explicar a dinâmica da criminalidade letal no estado de São Paulo, uma vez que a dinâmica temporal comum entre os grupos de controle e de tratamento (t) evidenciou que a criminalidade aumentaria com o passar do tempo, afinal o coeficiente subjacente ao tempo foi positivo e estatisticamente significativo. Além disso, o ajuste da regressão por efeitos fixos é melhorado com a inclusão da variável INFOCRIM. Essa afirmação pode ser confirmada por meio da comparação entre os critérios de informação BIC da regressão por efeitos fixos e do teste de sensibilidade.

Entretanto, conforme apontado por Matta e Andrade (2005) e Santos (2013), as interações espaciais podem ser relevantes para explicar o efeito médio de uma política de segurança pública. Nesse contexto, Justus e Kassouf (2008) argumentam que os estudos empíricos que analisam a criminalidade em nível regional devem considerar os aspectos espaciais dos dados, pois é possível que regiões próximas apresentem efeitos espaciais.

Ademais, conforme encontrado na seção 3.6, há evidências de um padrão espacial da atividade criminosa no estado de São Paulo.

Levando em consideração essas observações e, tendo em vista que, na presença de dependência espacial, o parâmetro de interesse \emptyset estimado por DID estará sub ou sobreidentificado, foi realizada a verificação da autocorrelação espacial por intermédio do teste I de Moran nos resíduos da regressão DID estimada por efeitos fixos (parte inferior da Tabela 4.1). A partir da adoção de diversas matrizes espaciais (k vizinhos mais próximos, sendo $k=1, \dots, 8, 10, 15$ e 20 e matriz de distância inversa), rejeita-se a hipótese nula de aleatoriedade espacial dos resíduos para a maioria das matrizes testadas. Seguindo o procedimento de Baumont (2004), a matriz de ponderação espacial que melhor captou a dependência espacial da criminalidade (maior valor da estatística I de Moran significativa para os resíduos da regressão DID) foi a matriz dos vinte vizinhos mais próximos ($k20$)²⁴.

Diante desse resultado, de fato, a dependência espacial parece ser importante fator para explicar a trajetória da taxa de homicídios por cem mil habitantes no estado de São Paulo. Com isso, não é possível garantir a hipótese SUTVA e, portanto, o parâmetro de interesse (\emptyset) não está identificado corretamente no modelo DID. Sendo assim, a metodologia SDID é a mais adequada para avaliar o impacto do INFOCRIM.

Os modelos SDID foram estimados pelo procedimento proposto por Elhorst (2003) com erros-padrão robustos (matriz de variância-covariância de White). A matriz de ponderação espacial adotada nas estimações é a matriz baseada no critério de vizinhança considerando os vinte vizinhos mais próximos ($k20$), conforme definição do procedimento de Baumont (2004).

Os resultados da taxonomia dos modelos SDID podem ser visualizados na Tabela 4.3. As colunas de 2 a 5 referem-se às estimativas do modelo de defasagem espacial (SDID-SAR), enquanto, na sexta coluna, estão reportados os resultados do modelo de erro espacial (SDID-SEM); nas colunas 7, 8 e 9 estão as estimações dos modelos regressivo cruzado espacial (SDID-SLX), de Durbin espacial (SDID-SDM) e de erro espacial e transbordamentos espaciais (SDID-SDEM), respectivamente e, na última coluna, encontram-se os resultados para o modelo de defasagem espacial com erro autorregressivo espacial (SDID-SAC).

Como discorrido na seção 3.4, após a estimação dos modelos SDID por efeitos fixos, deve-se escolher, como melhor modelo, aquele que obtiver o melhor ajuste apontado pelo valor do critério de informação BIC. Diante disso, dentre todos os modelos estimados, considerando

²⁴ O resultado do procedimento de Baumont (2004) para a escolha da melhor matriz de ponderação espacial encontra-se no Apêndice D.

os 615 municípios paulistas, o modelo SDID que melhor se ajusta aos dados é o modelo de defasagem espacial (SDID-SAR).

Com o objetivo de confirmar se a matriz $k20$ seria a mais adequada para estimar os modelos SDID, também foi realizado o procedimento de Stakhovych e Bijmolt (2009). O procedimento proposto pelos autores para a escolha da matriz de ponderação espacial consiste em sete passos e, de forma resumida, deve-se estimar os modelos SDID com diferentes matrizes de pesos espaciais e escolher aquela que obtiver o melhor critério de ajuste. Entre os modelos SDID estimados com diversas matrizes de ponderação espacial ($k = 1, \dots, 8, 10, 15, 20$ e matriz de distância inversa), o modelo SDID-SAR estimado com a matriz $k20$ continuou apresentando o melhor ajuste (menor critério de informação BIC).

Tabela 4.2 – Resultados das estimações dos modelos de diferenças em diferenças espacial, SDID

(continua)

Crime	SAR	SEM	SDM	SLX	SDEM	SAC
Tempo (t)	6,475** (3,076)	7,265** (3,452)	-14,952 (9,496)	-16,663 (13,691)	-15,098 (10,621)	6,064 (4,836)
Infocrim	-3,700* (2,032)	-5,023** (2,174)	-0,936 (2,201)	-0,891 (3,185)	-1,054 (2,206)	-3,482 (2,715)
Densidade (Dens)	-0,025*** (0,004)	-0,026*** (0,004)	-0,020*** (0,004)	-0,020*** (0,006)	-0,020*** (0,004)	-0,025*** (0,008)
Renda (Renda)	0,010 (0,007)	0,013* (0,008)	0,011 (0,008)	0,011 (0,011)	0,011 (0,008)	0,009 (0,010)
Desemprego (Des)	0,309** (0,144)	0,448*** (0,150)	0,351** (0,152)	0,341 (0,221)	0,343** (0,152)	0,268 (0,249)
Educação (Educ)	-0,539*** (0,152)	-0,624*** (0,155)	-0,585*** (0,156)	-0,575** (0,226)	-0,583*** (0,156)	-0,511** (0,215)
Sem religião (Srel)	23,908 (17,969)	22,275 (18,224)	16,215 (18,026)	15,318 (26,081)	16,193 (18,028)	24,026 (41,660)
Homens Jovens (Masc)	-18,135 (25,255)	-16,065 (25,801)	-12,471 (25,620)	-11,170 (37,069)	-11,162 (25,585)	-18,646 (30,453)
Lar uniparental (Chefe)	0,109 (0,127)	0,047 (0,129)	0,128 (0,128)	0,128 (0,185)	0,133 (0,128)	0,126 (0,205)
Bolsa Família (BF)	-44,572 (28,289)	-49,674 (32,919)	-41,397 (37,000)	-39,790 (53,539)	-42,775 (36,856)	-43,181 (60,314)
Desigualdade (r20)	-0,611** (0,282)	-0,706** (0,300)	-0,514* (0,309)	-0,499 (0,447)	-0,523* (0,308)	-0,579 (0,447)
Desarmamento (Saf)	-1,221 (1,828)	-1,115 (1,844)	-0,665 (1,831)	-0,651 (2,650)	-0,656 (1,837)	-1,282 (2,721)
Guarda Municipal (Garma)	-15,355*** (5,097)	-16,314*** (5,143)	-15,580*** (5,108)	-15,948** (7,389)	-16,032*** (5,114)	-15,064** (7,502)
WInfocrim			-15,102* (7,920)	-0,895 (0,565)	-0,936** (0,426)	
Wdens			-0,044*** (0,017)	-0,003** (0,001)	-0,003*** (0,001)	
Wrenda			0,017 (0,022)	0,001 (0,002)	0,001 (0,001)	

Tabela 4.3 – Resultados das estimações dos modelos de diferenças em diferenças espacial, SDID

(conclusão)

Crime	SAR	SEM	SDM	SLX	SDEM	SAC
Wdes			-1,404*** (0,509)	-0,074** (0,037)	-0,077*** (0,028)	
Weduc			0,555 (0,636)	0,032 (0,046)	0,023 (0,035)	
Wsrel			40,713 (76,580)	2,163 (5,541)	2,512 (4,222)	
Wmasc			26,391 (94,385)	1,811 (6,821)	1,269 (5,284)	
Wchefe			1,141* (0,616)	0,059 (0,045)	0,063* (0,034)	
WBF			4,513 (80,879)	0,268 (5,853)	-0,336 (4,485)	
Wr20			0,052 (0,950)	0,006 (0,069)	-0,002 (0,053)	
Wsaf			-2,700 (9,656)	-0,129 (0,699)	-0,175 (0,529)	
Wgamma			-4,167 (20,796)	-0,488 (1,492)	-0,754 (1,136)	
ρ	0,407*** (0,057)		0,180** (0,087)			0,453*** (0,084)
λ		0,429*** (0,079)			0,181** (0,091)	-0,137 (0,150)
Constante				28,688 (31,997)		
R ²	0,244	0,221	0,266	0,266	0,266	0,245
BIC	8878,372	8900,099	8934,809	8931,73	8935,044	8884,592
Número de observações	1230	1230	1230	1230	1230	1230

Fonte: Elaboração Própria.

Notas: Erros-padrão robustos entre parênteses;

*** p-valor<0,01, ** p-valor<0,05, * p-valor<0,1;

A estimação dos modelos espaciais por MV, conforme aconselhado por Elhorst (2003), foi implementada no *software* econométrico Stata 12 por meio do comando *xsmle* de Belotti *et al.* (2013).

Com a especificação SDID-SAR, o cálculo do efeito médio do tratamento (ATE) deve levar em consideração a transformação espacial - $(I - W\rho)^{-1}\theta$ - demonstrando que as interações espaciais entre todos os municípios influenciam o efeito médio do INFOCRIM. A decomposição do efeito total em efeitos direto e indireto do modelo SDID-SAR está reportada na Tabela 4.4.

Ao flexibilizar a hipótese SUTVA, considerando as interações espaciais no modelo SDID-SAR, é possível afirmar que o impacto do INFOCRIM encontrava-se sobreidentificado, uma vez que, ao invés de um efeito redutor no montante de 7,986, conforme estimação DID por efeitos fixos, o efeito médio do INFOCRIM aponta para um efeito de dissuasão total de 6,183 na taxa de homicídios por cem mil habitantes nos municípios do estado de São Paulo.

Tabela 4.3 – Decomposição do efeito total do modelo SDID-SAR em efeitos direto e indireto.

Variáveis	Efeito Total	Efeito Direto	Efeito Indireto
Tempo (t)	10,774** (5,172)	6,448** (3,053)	4,326* (2,293)
Infocrim	-6,183* (3,322)	-3,745* (2,036)	-2,437* (1,376)
Densidade (Dens)	-0,042*** (0,007)	-0,025*** (0,004)	-0,017*** (0,004)
Renda (Renda)	0,017 (0,012)	0,010 (0,007)	0,007 (0,005)
Desemprego (Des)	0,526** (0,239)	0,317** (0,146)	0,209** (0,102)
Educação (Educ)	-0,901*** (0,256)	-0,539*** (0,148)	-0,362*** (0,128)
Sem religião (Srel)	41,837 (29,594)	25,026 (17,479)	16,811 (12,726)
Homens Jovens (Masc)	-33,249 (43,299)	-19,812 (25,508)	-13,437 (18,317)
Lar uniparental (Chefe)	0,188 (0,217)	0,111 (0,126)	0,077 (0,093)
Bolsa Família (BF)	-74,031 (48,864)	-44,353 (29,159)	-29,677 (20,721)
Desigualdade (r20)	-1,032** (0,492)	-0,619** (0,291)	-0,413* (0,217)
Desarmamento (Saf)	-2,061 (2,984)	-1,243 (1,769)	-0,818 (1,242)
Guarda Municipal (Garma)	-25,942*** (9,597)	-15,553*** (5,641)	-10,390*** (4,440)

Fonte: Elaboração Própria.

Notas: Erros-padrão robustos entre parênteses;

*** p-valor<0,01, ** p-valor<0,05, * p-valor<0,10;

A análise dos impactos diretos revela que o INFOCRIM reduziu em 3,745 a taxa de homicídios por cem mil habitantes nos municípios onde esse sistema de

georreferenciamento foi implementado. Em relação aos efeitos indiretos, é possível afirmar que o Programa possui um efeito dissuasor indireto, tendo em vista que pode ser creditado ao INFOCRIM uma redução média de 2,437 na taxa de crimes letais dos demais municípios paulistas. Portanto, a interação espacial existente entre os agentes potencializou o efeito médio da política. Diante desse resultado, pode-se creditar a queda da criminalidade observada na década passada nos municípios paulistas ao efeito de dissuasão decorrente da implantação do INFOCRIM.

Quanto aos efeitos estimados dos determinantes da criminalidade do modelo SDID-SAR, pode-se observar que a taxa de homicídios é impactada positivamente pela taxa de desemprego (*des*). Essa relação pode ser entendida em virtude de uma maior taxa de desemprego reduzir o custo de oportunidade dos indivíduos de cometerem crime, o que, por sua vez, aumenta os índices de criminalidade letal. Sendo assim, as condições do mercado de trabalho formal parecem ser importantes fatores para explicar a criminalidade. Evidências da relação positiva entre crime e desemprego também foram encontrados por Pereira e Fernandez-Carrera (2000) para a região da grande São Paulo e Justus e Kassouf (2013) para a cidade de São Paulo.

Por outro lado, o acúmulo de capital humano decorrente da educação aumenta as oportunidades de obtenção de maiores retornos esperados no mercado de trabalho legal, reduzindo as taxas de homicídios. Portanto, uma maior parcela da população com, pelo menos, o nível médio de escolaridade (*educ*) possui maiores custos de oportunidade para a prática do crime de tal modo a reduzir a oferta de crime nos municípios paulistas. Evidências empíricas da relação negativa entre crime e educação no Brasil também foram encontrados por Araújo Júnior e Fajnzylber (2001) e Teixeira (2011).

No que se refere à desigualdade de renda, o retorno da atividade criminosa dos indivíduos pobres, que possuem perspectivas de baixos retornos no mercado legal, tende a aumentar a oferta do crime ao se depararem com indivíduos de alta renda (KELLY, 2000). No caso dos municípios paulistas, cabe observar que a desigualdade social reduziu consideravelmente na primeira década dos anos 2000. A melhoria do cenário de distribuição social contribuiu para a redução dos índices de criminalidade violenta nos municípios do estado de São Paulo, indicada pelo coeficiente negativo e estatisticamente significativo da variável r_{20} . Em outras palavras, como a distribuição de renda tornou-se mais equitativa no período analisado, o coeficiente negativo da variável r_{20} pode ser entendido no sentido de que uma menor desigualdade de renda reduziu a criminalidade nos municípios paulistas.

Já o impacto da densidade demográfica sobre a criminalidade apresentou sinal contrário ao esperado, uma vez que o adensamento populacional parece ter reduzido a criminalidade violenta no estado de São Paulo no período analisado. Esse resultado pode ser entendido pelo fato de que as taxas de homicídios reduziram de forma mais acelerada nos municípios com alta densidade demográfica (WAISELFISZ, 2011).

Controlando para a iniciativa municipal na esfera de segurança pública, pode-se perceber que a existência de guarda municipal com porte de arma de fogo (*garma*) parece ter reduzido a criminalidade nos municípios de São Paulo. Apesar de a guarda municipal ter como principal atribuição a proteção do patrimônio do município, a redução da criminalidade letal observada pode ser decorrente de seu efeito dissuasor indireto.

Há de se observar que variáveis como renda, percentagem de jovens do sexo masculino, percentagem da população sem religião e a desorganização familiar não se apresentaram como fatores relevantes para explicar a trajetória de criminalidade nos municípios do estado de São Paulo no período considerado.

Resultados interessantes referem-se às *proxies* utilizadas para o Estatuto do Desarmamento e para o programa de transferência de renda do Governo Federal, o Bolsa Família. Apesar de a literatura nacional apontar efeito de dissuasão na criminalidade decorrente dessas políticas, tais iniciativas não foram importantes para explicar a dinâmica recente da criminalidade letal para os 615 municípios paulistas. Ainda que a *proxy* utilizada para captar o efeito do Estatuto do Desarmamento (*saf*) tenha apresentado sinal negativo em consonância com os resultados encontrados por Cerqueira e Mello (2012) e Justus e Kassouf (2012, 2013), a variável não foi estatisticamente significativa.

Quanto à *proxy* para o Bolsa Família, também não foram encontrados efeitos significativos do programa de transferência de renda sobre as taxas de crimes contra a pessoa (homicídios dolosos e culposos) em conformidade com os resultados do estudo de Loureiro (2012). Por outro lado, o estudo de Chioda *et al.* (2015) sugere um efeito dissuasor sobre a criminalidade ao melhorar as condições socioeconômicas dos indivíduos para a cidade de São Paulo. Ainda, vale ressaltar que o estudo de Loureiro (2012) mostra que o Bolsa Família foi importante para reduzir os crimes contra o patrimônio para os estados brasileiros no período de 2001 a 2008.

Diante dos resultados encontrados por essa tese, é possível afirmar que a política de segurança pública implementada pelo Governo do Estado de São Paulo foi exitosa em reduzir as taxas de homicídios por cem mil habitantes no período analisado. Em termos

de vidas poupadas, pode-se inferir que o INFOCRIM evitou 2.546 homicídios no estado de São Paulo desde sua implementação até o ano de 2010.

5 CONSIDERAÇÕES FINAIS E IMPLICAÇÕES POLÍTICAS

A notória reversão dos índices de criminalidade letal nos municípios paulistas nos anos recentes tem motivado o interesse dos pesquisadores em verificar as possíveis causas e/ou avaliar as políticas de segurança pública implementadas que possam ter contribuído para a redução das taxas de homicídios. Esse interesse é justificado pelo fato de o estado de São Paulo ter se tornado um caso de sucesso reconhecido nacional e internacionalmente no combate à criminalidade, comparável a Bogotá e Nova Iorque.

Nesse contexto, a contribuição da presente tese para a literatura de economia do crime, mais especificamente para a avaliação de política, consistiu em analisar o impacto do INFOCRIM sobre a criminalidade nos municípios paulistas, controlando diversos determinantes do crime e outras políticas de segurança pública implementadas no período. Ademais, apesar das evidências encontradas na literatura acerca da possibilidade de migração do crime das regiões tratadas para regiões vizinhas após a implementação de políticas de segurança pública, as avaliações de tais políticas desconsideraram essas possíveis interações espaciais existentes. Para resolver esse problema, a presente tese propôs o uso pioneiro da metodologia de diferenças em diferenças espacial para identificar corretamente o impacto do INFOCRIM sobre a criminalidade letal em São Paulo.

De modo a evidenciar a relevância da inclusão do espaço na avaliação de tratamento, a AEDE realizada confirmou a presença de dependência espacial global uni e bivariada para a taxa de homicídios por cem mil habitantes e para os demais determinantes no caso dos 615 municípios paulistas contemplados na amostra no período compreendido entre os anos de 2000 e 2010, em um padrão de concentração. Tal padrão de aglomeração da atividade criminosa é mantido com a identificação de *clusters* espaciais de alta criminalidade nos municípios da RMSP no início do período analisado. Já no ano de 2010, o número de municípios pertencentes ao *cluster* de alta criminalidade reduziu consideravelmente, ao passo que mais municípios passaram a fazer parte do *cluster* de baixos índices de homicídios. Esses fatos sugeriram que a omissão da influência da autocorrelação espacial levaria a má especificação do modelo de diferenças em diferenças, causando viés no efeito causal estimado do INFOCRIM sobre a criminalidade.

A análise econométrica realizada partiu de um modelo DID estimado por efeitos fixos, confirmado pelo teste de Hausman (1978) como o mais apropriado para tratar os efeitos não observados, no qual foram encontrados indícios do impacto do INFOCRIM sobre a redução da criminalidade. Entretanto, o teste de autocorrelação espacial *I* de Moran nos resíduos da regressão DID confirmou que, de fato, as interações espaciais são fatores importantes para explicar a criminalidade nos municípios paulistas. Diante disso, os modelos SDID foram estimados com o intuito de identificar corretamente o efeito causal do INFOCRIM sobre as taxas de homicídios no estado de São Paulo. O modelo de defasagem espacial – SDID-SAR – foi o que se mostrou mais adequado para avaliar a política de segurança pública implementada. A migração da criminalidade entre os municípios paulistas foi confirmada em virtude da significância do parâmetro de defasagem espacial (ρ).

Os resultados obtidos por essa tese, no que se refere aos determinantes da criminalidade, comprovaram a relação positiva entre o crime e o desemprego, sugerindo que uma maior taxa de desemprego reduz o retorno esperado no mercado de trabalho formal, assim como o custo de oportunidade dos indivíduos de cometerem crime. Por outro lado, a melhora na educação contribuiu para a queda da criminalidade, uma vez que o maior nível de capital humano dos indivíduos aumenta suas oportunidades de obtenção de maiores retornos esperados no mercado de trabalho legal. Diante disso, as condições do mercado de trabalho formal devem ser melhoradas (redução do desemprego e políticas de qualificação da população, por exemplo) com o intuito de, por um lado, aumentar o custo de oportunidade da prática criminosa e, por outro, aumentar o rendimento esperado na atividade legal.

A melhoria do cenário de distribuição social ocorrida nos municípios paulistas na década passada também contribuiu para a redução dos índices de criminalidade violenta. O adensamento populacional, por sua vez, apresentou sinal contrário ao esperado. Uma das possíveis explicações estaria na observação de que as taxas de homicídios reduziram de forma mais acelerada nos municípios com alta densidade demográfica.

Controlando outras ações de segurança pública, foi observado o efeito dissuasor indireto sobre as taxas de homicídio decorrente da existência de guarda municipal com porte de arma de fogo. Já as variáveis renda, percentagem de jovens do sexo masculino, percentagem da população sem religião, a desorganização familiar, o Estatuto do Desarmamento e o programa de transferência de renda do Governo Federal - Bolsa

Família - não foram significativos para explicar a dinâmica recente da criminalidade letal nos municípios do estado de São Paulo.

Vale salientar que, como em qualquer avaliação de tratamento, os determinantes da criminalidade apontados pela teoria econômica do crime foram incluídos nas regressões de modo a estabelecer o grupo de controle como um bom contrafactual e assim, garantir as condições necessárias para identificar corretamente o impacto do INFOCRIM sobre a criminalidade.

Sendo assim, a constatação mais interessante desta tese deve-se ao efeito de dissuasão decorrente da implantação do INFOCRIM. Essa política específica de combate ao crime na qual o aparato policial foi melhorado com a introdução de novas tecnologias contribuiu para a reversão dos índices de criminalidade violenta nos municípios paulistas, sendo responsável por uma redução média de 6,183 sobre a taxa de homicídios por cem mil habitantes na década passada. No contexto da teoria econômica do crime, é possível afirmar que a aplicação mais rigorosa e imediata da lei pelo Governo do estado de São Paulo tenha alterado o comportamento do potencial criminoso no sentido de reduzir a oferta de crime em virtude da percepção de aumento da probabilidade de prisão e, por conseguinte, da probabilidade de fracasso do ato criminoso.

O efeito dissuasor direto na criminalidade letal nos 67 municípios que implementaram o INFOCRIM foi de 3,745, enquanto a redução média de 2,437 na taxa de homicídios dos demais municípios paulistas é decorrente do efeito dissuasor indireto do INFOCRIM. Portanto, a interação espacial existente entre os agentes potencializou o efeito médio da política de segurança pública do Estado de São Paulo. Isso constitui uma informação útil para os *policy makers* no que se refere ao entendimento dos resultados das políticas públicas implementadas, ante aos seus resultados esperados, de tal maneira a subsidiar as decisões governamentais.

Diante do ocorrido, a política de segurança pública do Estado de São Paulo pode ser considerada exitosa, tendo em vista que, ao reduzir as taxas de homicídios por cem mil habitantes no período analisado, o Estado conseguiu minimizar os prejuízos sociais decorrentes da criminalidade. Em termos de vidas poupadas, o INFOCRIM evitou 2.546 homicídios no estado de São Paulo desde sua implementação até o ano de 2010.

Como implicação política, os governos federal, estadual e municipal deveriam atuar de forma conjunta no combate à criminalidade com políticas de segurança pública que atuassem na aplicação rigorosa da lei. Tais políticas deveriam aumentar a probabilidade de prisão e detenção, bem como a efetividade do judiciário e, para tanto,

poderiam ser baseadas na adoção de tecnologias disponíveis para a melhora do desempenho da polícia e para o rápido reconhecimento do criminoso. Sendo assim, os demais estados do País deveriam adotar programas de segurança pública nos moldes do sistema de georreferenciamento das informações criminais, o INFOCRIM, para combate e prevenção da criminalidade.

Como agenda de pesquisa futura, seria interessante avaliar diferentes tipos de políticas de segurança pública implementadas nos demais estados do País, como as Unidades de Polícia Pacificadora – UPPs – no Rio de Janeiro, considerando a possível migração da criminalidade para regiões vizinhas.

REFERÊNCIAS

- ABADIE, A.; DIAMOND, A.; HAINMUELLER, J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California Tobacco Control Program, **Journal of the American Statistical Association**, American Statistical Association, vol. 105(490), pages 493-505, 2010.
- ALMEIDA, E. **Econometria Espacial Aplicada**, 1ª ed., Campinas: Editora Alínea, 2012.
- ALMEIDA, E. S. Impact Evaluation of Local Public Policies. **In: 61st annual North American Meeting of Regional Science Association International, 2014**, Washington, DC. Annals of 61st annual North American Meeting of Regional Science Association International, 2014.
- ALMEIDA, E. S., HADDAD, E. A.; HEWINGS, G. J. D. The spatial pattern of crime in Minas Gerais: An exploratory analysis. **Economia Aplicada**, vol. 9, n. 1, p. 39–55, 2005.
- ANGRIST, D. J.; KRUEGER, A. B.; Empirical strategies in labor economics. In: Ashenfelter, O; Card, D. (ed.), **Handbook of Labor Economics**, Elsevier, 1ª edição, p. 1277-1366, 1999.
- ANSELIN, L. **Spatial econometrics: methods and models**. Kluwer Academic, Boston, 1988.
- ANSELIN, L.; BERA, A. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah A. e Giles D. E. (eds.) **Handbook of applied economic statistics**, Marcel Dekker, Nova York, p. 237-289, 1998.
- ARAUJO, A.F.V.; RAMOS, F.S. Estimação da Perda de Bem-Estar Causada pela Criminalidade: O Caso da Cidade de João Pessoa – PB. **Economia**, v.10, n.3, p.577–607, 2009.
- ARAUJO JUNIOR, A.; FAJNZYLBER, P. Crime e economia: Um estudo das microregiões mineiras. **Revista Econômica do Nordeste**, 31(especial):630-659, 2000.
- ARAUJO JUNIOR, A.; FAJNZYLBER, P. **O que causa a criminalidade violenta no Brasil? Uma análise a partir do modelo econômico do crime: 1981 a 1996**. Belo Horizonte: Universidade Federal de Minas Gerais, CEDEPLAR, 88p. Texto de Discussão 162, 2001.
- ARRAIZ, I., DRUKKER, D. M., KELEJIAN, H. H., PRUCHA, I. R. A Spatial Cliff-Ord-type Model with Heteroskedastic Innovations: Small and Large Sample Results. **Journal of Regional Science**, vol. 50, n. 2, p. 592-614, 2010.
- BARTLEY, W. A.; COHEN, M. A. The effect of concealed weapons laws: an extreme bound analysis. **Economic Inquiry**, v. 36, n. 2, 1998.
- BAUMONT, C. **Spatial effects in housing price models: do house prices capitalize urban development policies in the agglomeration of Dijon (1999)?** Université de Bourgogne, 2004.
- BEATO FILHO, C.C. Políticas públicas de segurança e a questão policial. **São Paulo em Perspectiva**, São Paulo, v. 13, n. 4, out./dez. 1999

- BECCARIA, C. **On Crimes and Punishments**, Bobbs-Merril, Indianapolis, (1963 [1764])
- BECKER, Gary. S. Crime and Punishment: An Economic Approach. **Journal of Political Economy**, v. 76, 169-217, 1968.
- BELL, K. P.; BOCKSTAEL, N. E. Applying the generalized moments estimation approach to spatial problems involving microlevel data. **The Review of Economics and Statistics**, v. 82, n.1, p. 72-82, 2000.
- BELOTTI, F.; HUGHES, G.; MORTARI, A.P. **XSMLE - A Command to Estimate Spatial Panel Models in Stata**. 2013 German Stata Users Group Meeting, Potsdam, 2013.
- BENTHAM, J. **An Introduction to the Principles of Morals and Legislation**, Athlone Press, London, (1970 [1789]).
- BENOIT, J.P.; OSBORNE, J. M. Crime, Punishment, and Social Expenditure. **Journal of Institutional and Theoretical Economics**, vol. 151, p. 326–347, 1995.
- BIDERMAN, C.; MELLO, J. M. P.; SCHNEIDER, A.A. Dry Laws and Homicides: Evidence from the São Paulo Metropolitan Area. **Economic Journal**, vol. 120, p. 157-182, 2010.
- BLOCK, M.; HEINEKE, J. A Labour Theoretical Analysis of Criminal Choice. **American Economic Review**, vol. 65, p. 314-325, 1975.
- BLUMSTEIN, A. Youth Violence, Guns, and the Illicit-Drug Industry. **The Journal of Criminal Law and Criminology** 86 (1): 10-36, 1995.
- BLUNDELL, R. W.; MACURDY, T. E. Labor Supply: A Review of Alternative Approaches. In: Ashenfelter, O; Card, D. (ed.), **Handbook of Labor Economics**, Elsevier, 1ª edição, p. 1559-1695, 1999.
- BOURGUIGNON, F.; MORRISON, A. **Measuring the Social Cost of Crime and Violence: Methodological Issues**. LACEA, 2000.
- BOWLING, B. **Violent Racism: Victimisation, Policing and Social Context**. Revised Edition. Oxford: Oxford University Press, 1999.
- BRAGA A. A.; KENNEDY D. M.; WARING E. J.; PIEHL, A. M. Problem-oriented policing, deterrence, and youth violence: An evaluation of Boston's Operation Ceasefire. **Journal of Research in Crime and Delinquency**, 38, 195-225, 2001.
- BRAGA, A. A. **Effect of hot spots policing on crime**. Campbell Collaboration Systematic Review Final Report, 2007.
- BRAGA, A. A.; BOND, B.J. Policing Crime and Disorder Hot Spots: A randomized controlled trial. **Criminology**, 46(3), 577-607, 2008.
- BRAGA, A.A.; WEISBURD, D.L. **Police Innovation and Crime Prevention: Lessons Learned from Police Research over the Past 20 Years**. Paper presented at the National Institute of Justice (NIJ) Policing Research Workshop: Planning for the Future, Washington, DC, November 28-29, 2006.
- BRATTON, W.J. **Cutting Crime and Restoring Order: What America Can Learn from New York's Finest**. The Heritage Lectures, n. 573, 1996.
- BRONARS, S. G.; LOTT JUNIOR, J. R. Criminal deterrence, geographic spillovers, and the right to carry concealed handguns. **American Economic Review**, May 1998.

- BUVINIC, M., MORRISON, A. **La violencia como obstáculo para el desarrollo.** Washington, D.C. Banco Interamericano de Desarrollo, Oficina del Economista Principal, 2001.
- CAMERON, A. C., TRIVEDI, P. K. **Microeconomics: Methods and Applications.** Cambridge University Press, Cambridge, 2005.
- CARDOSO, F.L.M.G; CECCHETTO, F.R; CORRÊA, J.S; SOUZA, T.S. Homicídios no Rio de Janeiro, Brasil: uma análise da violência letal. **Ciência & Saúde Coletiva**, vol. 21, n. 4, p. 1277-1288, 2016.
- CARVALHO, A.X.Y.; CERQUEIRA, D.R.C.; RODRIQUES, R.I; LOBÃO, W.J.A. Custos das mortes por causa externas no Brasil, In: **Anais do XXXVI Encontro Nacional de Economia**, ANPEC - Associação Nacional dos Centros de Pós-Graduação em Economia, 2008.
- CERQUEIRA, D.; LOBÃO, W. **Determinantes da criminalidade: Uma resenha dos modelos teóricos e resultados empíricos.** Texto para Discussão 956, IPEA, 2003.
- CERQUEIRA, D. R. C. **Causas e Consequências do Crime no Brasil.** 2010, 168 p. Tese (Doutorado em Economia) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2010.
- CERQUEIRA, D.; MELLO, J. M. **Menos Armas, Menos Crimes.** Brasília: IPEA – Texto para Discussão nº. 1721, mar/2012.
- CERQUEIRA, D.; CARVALHO, A.; LOBÃO, W.; RODRIGUES, R. **Análise dos custos e consequências da violência no Brasil.** Texto para discussão IPEA no. 1284, 2007.
- CHAGAS, A.L.S.; AZZONI, C. R.; ALMEIDA, A.N. Sugarcane burning and human health: a spatial difference-in-difference analysis. In: **36 Congresso da SBE, 2014**, Natal. Anais do 36 Congresso da SBE, 2014.
- CHIODA, L.; MELLO, J. M. P.; SOARES, R. R. Spillovers from conditional cash transfer programs: *Bolsa Família* and Crime in urban Brazil. **Economics of Education Review**, In Press, 2015, <http://dx.doi.org/10.1016/j.econedurev.2015.04.005>.
- COHEN, L.; FELSON, M. Social change and crime rate trends: a routine approach. **American Sociological Review**, 44, p.588-608, 1979.
- CORMAN, H.; MOCAN, N. Carrots, Sticks, and Broken Windows. **Journal of Law and Economics**, vol. 48, p. 235-66, 2005.
- DIAS, C.C.N. **Da pulverização ao monopólio da violência: expansão e consolidação do Primeiro Comando da Capital (PCC) no sistema carcerário paulista.** 2011. 386f. Tese (Doutorado em Sociologia) - Departamento de Sociologia, Universidade de São Paulo, São Paulo, 2011.
- DONOHUE, J.J.; LEVITT, S.D. The Impact of Legalized Abortion on Crime. **The Quarterly Journal of Economics**. Vol. CXVI, p. 379-420, 2001.
- DUBÉ, J., LEGROS, D., THÉRIAULT, M., DES ROSIERS, F. A spatial Difference-in-Differences estimator to evaluate the effect of change in public mass transit systems on house prices. **Transportation Research Part B**, vol. 64, p. 24–40, 2014.
- DUGGAN, M. More Guns, More Crime, **Journal of Political Economy**, v109, p. 1086-1114, 2001.

- ECK, J.; MAGUIRE, E. **Have Changes in Policing Reduced Violent Crime? An assessment of the evidence.** In *The Crime Drop in America*, (Eds.). Alfred Blumstein and Joel Wallman. New York: Cambridge University Press, 2000.
- EHRlich, I. **The supply of illegitimate activities.** Unpublished manuscript, New York: Columbia University, 1967.
- EHRlich, I. Participation in illegitimate activities: A theoretical and empirical investigation. **Journal of Political Economy**, vol. 81, p. 526-536, 1973.
- ELHORST, J. Specification and estimation of spatial panel data models. **International Regional Science Review**, Vol. 26, N. 3, p. 244–268, 2003.
- ELHORST, J. Applied Spatial Econometrics: Raising the Bar. **Spatial Economic Analysis**, vol. 5, n. 1, p. 9-28, 2010.
- ENDERS, W. **Applied Econometric Times Series**, 3rd Edition, 2010. 544p.
- FAJNZYLBER, P.; LEDERMAN, D.; LOAYZA, N. Crime and Victimization: an Economic Perspective. **Economia**, vol. 1, p. 219-302, 2000.
- FAJNZYLBER, P.; ARAUJO JUNIOR, A.F. Violência e criminalidade. In: LISBOA, M.B.; MENEZES FILHO, N.A. (Ed), **Microeconomia e Sociedade no Brasil**. Contra Capa: Rio de Janeiro, p. 333-395. 2001.
- FERREIRA, R.M.C. **Uma avaliação do programa das UPPs sobre o seu impacto nas taxas de criminalidade da cidade do Rio de Janeiro e da Baixada Fluminense.** 2013. 28 f. Monografia (Bacharelado em Economia) - Departamento de Economia, Pontifícia Universidade Católica, Rio de Janeiro, 2013.
- FLEISHER, B. M. The effect of unemployment on juvenile delinquency. **The Journal of Political Economy**, Vol. 71, n. 6, p. 543–555, 1963.
- FLEISHER, B. M. The effect of income on delinquency. **The American Economic Review**, vol. 61, n. 1, p. 118–137, 1966.
- FLOOD-PAGE, C.; CAMPBELL, S.; HARRINGTON, V.; MILLER, J. **Youth crime: findings from the 1998/99 youth lifestyles survey.** Home Office Research, Development and Statistics Directorate Crime and Criminal Justice Unit, United Kingdom, 2000. Home Office Research Study 209, 2000.
- FREEMAN, R.B. Crime and the Labor Market. In James, Q. Wilson (Ed.), **Crime and Public Policy**. San Francisco: Institute for Contemporary Studies, 1983.
- FREEMAN, R.B. Why do so many young American men commit crimes and what might we do about it? **Journal of Economic Perspectives**, vol. 10, p. 25-42, 1996.
- GLAESER, E.; SACERDOTE, B. Why is There More Crimes in Cities. **Journal of Political Economy**, v.107, n.6, p.225-258, 1999.
- GOERTZEL, T.; KAHN, T. The Great São Paulo Homicide Drop. **Homicide Studies**, vol. 13, n. 4, p. 398–410, 2009.
- GOLDSTEIN, H. Improving Policing: A Problem Oriented Approach, **Crime and Delinquency**, vol. 25, April, pp. 236-58, 1979.
- GRAHAM, J.; BOWLING, B. **Young people and crime.** United Kingdom: Great Britain Home Office. Research and statistics Dept., 1995.

- HARTUNG, G.C. **Ensaio em Demografia e Criminalidade**. 2009, 101 p. Tese (Doutorado em Economia) - Rio de Janeiro: EPGE, 2009.
- HAUSMAN, J. A. Specification tests in econometrics. **Econometrica**, vol. 46, n. 6, p. 1251–1272, 1978.
- HEATON, P. Does Religion Really Reduce Crime? **The Journal of Law and Economics**, vol. 9, n. 1, p. 147-172, 2006.
- HECKERT, M.; MENNIS, J. The economic impact of greening urban vacant land: a spatial difference-in-differences analysis, **Environment and Planning A**, vol. 44(12), p. 3010-3027, 2012.
- HIRSCHI, T.; STARK, R. Hellfire and Delinquency. **Social Problems**, vol. 17, p. 202–213, 1969.
- HULL, B.B. Religion Still Matters. **Journal of Economics**, vol. 26, n. 2, p. 35-48, 2000.
- IMROHOROGLU, A.; MERLO, A.; RUPERT, P. On the Political Economy of Income Redistribution and Crime. **International Economic Review**, vol. 41, n. 1, p. 1–25, 2000.
- JANG, H.; HOOVER, L.T.; JOO, H-J. An Evaluation of Compstat's Effect on Crime: The Fort Worth Experience. **Police Quarterly**, vol. 13, n. 4, p. 387-412, 2010
- JUSTUS, M.; KASSOUF, A. L. Uma investigação econômica da influência do mercado de drogas ilícitas sobre a criminalidade brasileira. *Revista EconomiA*, vol. 8, n. 2, p. 187–210, 2007.
- JUSTUS, M.; KASSOUF, A. L. Existe explicação econômica para a taxa de crimes contra a propriedade? **Economia Aplicada**, v. 12, n.1, 2008.
- JUSTUS, M.; KASSOUF, A L. Menos Desemprego, maior eficiência da polícia e redução na criminalidade em São Paulo. **In: 39 Encontro Nacional de Economia ANPEC, 2011**, Foz do Iguaçu. 39 Encontro Nacional de Economia ANPEC, 2011.
- JUSTUS, M.; KASSOUF, A L. Avaliação de Impacto do Estatuto do Desarmamento na Criminalidade: Uma abordagem de séries temporais aplicada à cidade de São Paulo. **Economic Analysis of Law Review**, vol. 3, nº 2, p. 307-322, 2012.
- JUSTUS, M.; KASSOUF, A. L. A cointegration analysis of crime, economic activity, and police performance in São Paulo city. **Journal of Applied Statistics**, v. 1, p. 1-23, 2013.
- JUSTUS, M.; SANTOS FILHO, J.I. Convergência das Taxas de Crimes no Território Brasileiro. **EconomiA**, Brasília (DF), v.12, n.1, p.131–147, jan/abr 2011.
- KAHN, T. The homicide drop in São Paulo, Brazil. Boston, 2007.
- KAHN, T.; Zanetic, A. Estudos criminológicos 4 – o papel dos municípios na segurança pública, 2005.
- KATZ, C. M.; WEBB, V. J.; SCHAEFER, D. R. An assessment of the impact of quality-of-life policing on crime and disorder. **Justice Quarterly**, vol. 18(4), p. 825-876, 2001.
- KELEJIAN, H. H.; PRUCHA, I. R. **A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances**. Mimeo., Department of Economics, University of Maryland, 1998.

- KELEJIAN, H. H.; PRUCHA, I. R. A generalized moments estimator for the autoregressive parameter in a spatial model. **International Economic Review**, vol. 40, n. 2, 1999.
- KELEJIAN, H. H.; PRUCHA, I. R. HAC estimation in a spatial framework. **Journal of Econometrics**, n. 140, p. 131-154, 2007.
- KELEJIAN, H. H.; PRUCHA, I. R. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. **Journal of Econometrics**, vol. 157, n. 1, p. 53–67, 2010.
- KELEJIAN, H. H.; ROBINSON, D. P. Spatial autocorrelation: A new computationally simple test with an application to per capita country police expenditures. **Regional Science and Urban Economics**, 22:317–333, 1992.
- KELLING, G. L.; SOUSA, W. H. **Do Police Matter? An Analysis of the Impact of New York City's Police Reforms**. Civic Report N^o. 22. New York, NY: Manhattan Institute, 2001.
- KELLY, M. Inequality and crime. **The Review of Economics and Statistics**, Vol. 82, n. 4, p. 530-539, 2000.
- KLECK, G. **Targeting guns: firearms and their control**. New York: Walter de Gruyter, Inc., 1997.
- KLECK, G. Measures of gun ownership levels for macro-level crime and violence research. **Journal of Research in Crime and Delinquency**, v. 41. n. 1, p. 3-36, fev. 2004.
- LEE, L-F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. **Econometrica**, vol. 72, n. 6, p. 1899-1925, 2004.
- LESAGE, J. P.; PACE, R. K. **Introduction to Spatial Econometrics**. CRC Press, Boca Raton, 2009.
- LEVITT, S.D. Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not. **Journal of Economic Perspectives**. Vol. 18, p. 163–190, 2004.
- LOTT, J.R. **More Guns, Less Crime: Understanding Crime and Gun Control Laws**, Chicago: The University of Chicago Press, 3rd Edition, p. 472, 2010.
- LOTT JUNIOR, J. R.; MUSTARD, D. B. Crime, deterrence, and right-to-carry concealed handguns. **Journal of Legal Studies**, v. 26, Jan. 1997.
- LOUREIRO, A. **Can Conditional Cash Transfers Reduce Crime? Evidence from Brazil**. Mimeo, 2012.
- MACHIN, S.; MARIE, O. Crime and Police Resources: The Street Crime Initiative, **Journal of the European Economic Association**, European Economic Association, vol. 9(4), pages 678-701, 08, 2011.
- MANSKI, C. F. Identification of endogenous social effects: the reflection problem. **Review of Economic Studies**, vol. 60, p. 531-542, 1993.
- MATTA, R.A.; ANDRADE, M.V. Avaliação Econômica Do Impacto Do Programa De Controle De Homicídios Fica Vivo. In: **Anais do XXXIII Encontro Nacional de Economia**, ANPEC - Associação Nacional dos Centros de Pós-graduação em Economia, 2005.

- MAZEROLLE, L., ROMBOUTS, S., MCBROOM, J. The impact of COMPSTAT on reported crime in Queensland. **Policing: An International Journal of Police Strategies and Management**, vol. 30, p. 237-256, 2007.
- MCINTOSH, C. Estimating Treatment Effects from Spatial Policy Experiments: An Application to Ugandan Microfinance, **The Review of Economics and Statistics**, vol. 90, issue 1, p. 15-28, 2008.
- MCMILLEN, D. Issues in spatial data analysis. **Journal of Regional Science**, vol. 50, n. 1, p. 119–141, 2010.
- MELLO, J. M. P.; SCHNEIDER, A. Assessing São Paulo's Large Drop in Homicides: the Role of Demography and Policy Interventions. In: Rafael Di Tella; Sebastian Edwards; Ernesto Schargrotsky. (Org.). **The Economics of Crime: Lessons for and from Latin America**. Chicago: University of Chicago Press and NBER, 2010, p. 207-235.
- MENDONÇA, M.; LOUREIRO, P.; SACHSIDA, A. Criminalidade e desigualdade social no Brasil. Rio de Janeiro: IPEA – Texto para Discussão nº 967, jul. 2003.
- MERTON, R.K. Social Structure and Anomie. **American Sociological Review** 3:672-682, 1938.
- MEYER, B. D.; Natural and Quasi-Experiments in Economics. **Journal of Business and Economic Statistics**, vol.13, p. 151-161. 1995.
- MOODY, C. E.; MARVELL, T. B. Guns and crime. **Southern Economic Journal**, 71(4), p. 720-736, 2002.
- MUSUMECI, B. S. e Soares, H. Base Nacional de Estatísticas Criminais – Análise e Avaliação. **Relatório de Consultoria para Secretaria Nacional de Segurança Pública**, 2005.
- NÓBREGA JUNIOR, J.M.P; ROCHA, E.C.; SANTOS, M. Os determinantes da criminalidade violenta no Brasil (1995-2004). In: **International Congress of the Latin American Studies Association**, 2009, Rio de Janeiro. Anais eletrônicos...Rio de Janeiro, 2009.
- OLIVEIRA, C. A. Criminalidade e o tamanho das cidades brasileiras: Um enfoque da economia do crime. In: **Anais do XXXIII Encontro Nacional de Economia**, Natal. ANPEC - Associação Nacional dos Centros de Pós Graduação em Economia, 2005.
- ORD, K. Estimation methods for models of spatial interaction. **Journal of the American Statistical Association**, vol. 70, n. 349, p. 120-126, 1975.
- PEIXOTO, B. T. **Determinantes da criminalidade no município de Belo Horizonte**. Master's thesis, Universidade Federal de Minas Gerais, CEDEPLAR, Belo Horizonte, 2003.
- PEIXOTO, BT; ANDRADE, MV; AZEVEDO, JP. Avaliação do Programa Fica Vivo no Município de Belo Horizonte. In: **XXXV Encontro Nacional de Economia**, 2007. Recife, 2007.
- PEIXOTO, BT; ANDRADE, MV; AZEVEDO, JP. **Avaliação econômica do Programa Fica Vivo: o caso piloto**. Texto para discussão nº 336, Belo Horizonte: UFMG / Cedeplar, 2008.

RATCLIFFE, J.; TANIGUCHI, T.; GROFF, E.; WOOD, J. The Philadelphia foot patrol experiment: A randomized controlled trial of police patrol effectiveness in violent crime hot spots. **Criminology**, Vol. 49, Issue 3, p. 795–831, 2011.

REY, J. S.; MONTOURI, B. D. US Regional Income Convergence: A Spatial Econometrics Perspective. **Regional Studies**, vol. 33.2, p. 143-156, 1999.

ROSENBAUM, P.; RUBIN, D.B. The Central Role of the Propensity Score in Observational Studies for Causal Effects, **Biometrika**, vol. 70, n°. 1, 1983.

ROSENFELD, R.; FORNANGO, R.; BAUMER, E. Did Ceasefire, Compstat, and Exile reduce homicide? **Criminology & Public Policy**, vol. 4, p. 419-450, 2005.

RUBIN, D. Assignment to a Treatment group on the basis of a Covariate. **Journal of Educational Statistics**. n. 2, p.1-26, 1977.

SABATES, R.; FEINSTEIN, L. Effects of Government Initiatives on Youth Crime. **Oxford Economic Papers**, 60 (3). pp. 462-483, 2008.

SANTOS, R.B. Effectiveness of Police in Reducing Crime and the Role of Crime Analysis. In: **Crime Analysis With Crime Mapping**. United States: SAGE Publications, 3rd Edition, Chapter 3, p. 352, 2013.

SARGAN, J.D. Wages and Prices in the United Kingdom: A Study in Econometric Methodology. In: **Econometric Analysis for National Economic Planning**, eds. P. E. Hart, G. Mills, and J. K. Whitaker, London: Butterworths, 1964.

SARTORIS, A. N. **Homicídios na cidade de São Paulo: uma análise de causalidade e autocorrelação espaço-tempo**. 2000. Tese (Doutorado em Economia) - Faculdade de Economia, Administração e Ciências Contábeis, Universidade de São Paulo, São Paulo, 2000.

SCORZAFAVE, L.G.; SOARES, M.K.; DORIGAN, T.A. Vale a pena pagar para desarmar? Uma avaliação do impacto da campanha de entrega voluntária de armas sobre as mortes com armas de fogo. **Estudos Econômicos**, São Paulo, vol.45, n.3, p.475-497, jul.-set. 2015

SCRIPILLITI, E.A.O. **Crime nos municípios paulistas: um estudo acerca dos condicionantes sócio-econômicos e demográficos que contribuem para maior criminalidade e quais os efeitos das diferentes políticas municipais de segurança para o combate à criminalidade**. 2006. Dissertação (Mestrado em Economia) – Faculdade de Economia, Administração e Ciências Contábeis, Universidade de São Paulo, São Paulo, 2006.

SHERMAN, L.; GOTTFREDSON, D.; MACKENZIE, D.; ECK, J.; REUTER, P.; BUSHWAY, S. **Preventing Crime: What Works, What Doesn't and What's Promising**. A Report to the United States Congress. Washington, DC: U.S. Department of Justice, 1997.

SHIKIDA, P.F.A.; ARAUJO JUNIOR, A.F.; SHIKIDA, D.C.; BORILL, S.P. Determinantes do comportamento criminoso: um estudo econométrico nas penitenciárias central, estadual e feminina de Piraquara (Paraná). **Pesquisa e Debate**, São Paulo, vol. 17, p. 125-148, 2006.

SKOGAN, W.G. **Disorder and Decline: Crime and the Spiral of Decay in American Neighborhoods**. Berkeley, California: University of California Press, 1990.

- SKOGAN, W.G. **Impact of policing on social disorder: Summary of findings.** Washington, DC: U.S. Department of Justice, Office of Justice Programs, 1992.
- SKOGAN, W.G.; FRYDL, K. **Fairness and Effectiveness in Policing: The Evidence.** Washington, D.C.: National Academies Press, 2004.
- SKOLNICK, J.; BAYLEY, D.H. **Policiamento Comunitário.** Tradução de Ana Luíza Amêndola Pinheiro. Organização: Nancy Cardia. São Paulo: Editora da Universidade de São Paulo, 125p. (Coleção Polícia e Sociedade, 6), 2002.
- SILVEIRA, A.; ASSUNÇÃO, R.; SILVA, B.; BEATO, C. Impacto do Programa Fica Vivo. **Revista de Saúde Pública**, vol. 44(3): 496–502, 2010.
- SILVEIRA NETO, R.M.; RATTON, J.L.; MENEZES, T.A.; MONTEIRO, C. Avaliação de Política Pública para Redução da Violência: o Caso do Programa Pacto pela Vida do Estado de Pernambuco. In: **41º Encontro Nacional de Economia, 2013.** Foz do Iguaçu, 2013.
- SILVERMAN, E.B. **NYPD fights crime: Innovative strategies in policing.** Boston: Northeastern University Press, 1999.
- SOUSA, W.H.; KELLING, G.L. Of “broken windows,” criminology, and criminal justice. In: D. WEISBURD; A. A. BRAGA (Eds.), **Police Innovation: Contrasting Perspectives** (pp.77-97). Cambridge: Cambridge University Press, 2006.
- SOUZA, M. F. M., MACINKO, J., ALENCAR, A. P., MALTA, D. C., MORAIS NETO, O. L. Reduction in firearm-related mortality and hospitalizations in Brazil after gun control. **Health Affairs**, v, 26, n, 2, pp. 575-584, 2007.
- SSP/SP – Secretaria de Segurança Pública do Estado de São Paulo. **São Paulo é destaque internacional na queda dos homicídios.** 2012. Disponível em: http://www.ssp.sp.gov.br/acoes/acoes_taxa-homicidios.aspx.
- SULIANO, D. C.; OLIVEIRA, J. L. Avaliação do Programa Ronda do Quarteirão na Região Metropolitana de Fortaleza (Ceará). **Revista Brasileira de Estudos Regionais e Urbanos**, vol. 07, n. 2, pp. 52-67, 2013.
- STAKHOVYCH, S.; BIJMOLT, T.H.A. Specification of spatial models: A simulation study on weights matrices. **Papers in Regional Science**, Volume 88, Issue 2, p. 389–408, 2009.
- STARK, R.; KENT, K.; DOYLE, D.P. Religion and Delinquency: The Ecology of a ‘Lost’ Relationship. **Journal of Research in Crime and Delinquency**, vol. 19, p. 4–24, 1982.
- STARK, R. Religion as Context: Hellfire and Delinquency One More Time. **Sociology of Religion**, vol. 57, p. 163–173, 1996.
- TEIXEIRA, E.C. **Dois ensaios acerca da relação entre criminalidade e educação.** 2011. Tese (Doutorado em Economia Aplicada) – Departamento de Economia, Administração e Sociologia, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, São Paulo, 2011.
- THORNBERRY, T. P. **Empirical support for interactional theory: a review of the literature.** In: Hawkins, J. D. (ed.). Some current theories of crime and deviance. New York: Cambridge University Press, p. 198- 235, 1996.

UNDP - United Nations Development Programme. Regional Human Development Report 2013-2014 - Citizen security with a human face: Evidence and Proposals for Latin America. New York, 2013. Disponível em: <http://www.latinamerica.undp.org/content/dam/rblac/docs/Research%20and%20Publications/IDH/UNDP-RBLAC-HDRLatAmExecSummary2013.pdf>

UNODC - United Nations Office on Drugs and Crime. United Nations Publication, Sales, n. E.11.XI.10, 2012.

VILLAVECES, A.; CUMMINGS P.; ESPITIA V.E.; KOEPESELL ,T.D.; MCKNIGHT B.; KELLERMANN, A.L. Effect of a ban on carrying firearms on homicide rates in two Colombian cities. **Journal of the American Medical Association**, v.283, n. 9, pp. 1205-1209, 2000.

WASELFISZ, J.J. **Mapa da Violência 2012 - os novos padrões da violência homicida no Brasil.** São Paulo, Instituto Sangari, 2011.

WEISBURD, D.; ECK, J. What can police do to reduce crime, disorder, and fear? **Annals of the American Academy of Political and Social Science**, vol. 593:42–65, 2004.

WEISBURD, D.; BRAGA, A.A. (Eds) **Police innovation: Contrasting perspectives.** New York: Cambridge Univ. Press, 2006.

WEISBURD, D.; TELEP, C.W.; HINKLE, J.C.; ECK, J.E. The Effects of Problem Oriented Policing on Crime and Disorder. **Campbell Systematic Reviews**, vol. 14, 2008.

WILLIS, G.D. **The Killing Consensus: Homicide detectives, police that kill and organized crime in São Paulo, Brazil.** 2013. 288s. Thesis (Ph. D. in Urban and Regional Studies) - Massachusetts Institute of Technology, Dept. of Urban Studies and Planning, 2013.

WILLIS, G.D. **The Killing Consensus: Police, Organized Crime and the Regulation of Life and Death in Urban Brazil.** Berkeley: University of California Press, 226 pages, 2015.

WILLIS, J. J., MASTROFSKI, S. D., WEISBURD, D. Making sense of COMPSTAT: A theory-based analysis of organizational change in three police departments. **Law and Society Review**. Vol. 41: 147–188, 2007.

WILSON, J. Q.; KELLING, G.L. **Broken Windows: the police and neighborhood safety.** Criminological perspectives essential readings. Vol. 249, p. 29-38, 1982.

WOOLDRIDGE, J. **Econometric analysis of cross section and panel data.** 2nd Edition, The MIT Press, 2010.

ZHANG, J. The Effect of Welfare Programs on Criminal Behavior: A Theoretical and Empirical Analysis,” *Economic Inquiry*, 35(1), 120–37.

APÊNDICES

Apêndice A – Lista dos municípios paulistas que implementaram o INFOCRIM, 2000 a 2010.

Município	Ano da implementação	Município	Ano da implementação
Americana	2008	Mairiporã	2003
Araçatuba	2008	Marília	2008
Arujá	2003	Mauá	2003
Barretos	2008	Mogi das Cruzes	2003
Barueri	2003	Mogi Guaçu	2008
Bauru	2005	Osasco	2003
Birigui	2008	Pindamonhangaba	2008
Biritiba-Mirim	2003	Piracicaba	2008
Caieiras	2003	Pirapora do Bom Jesus	2003
Cajamar	2003	Poá	2003
Campinas	2003	Presidente Prudente	2008
Caraguatatuba	2008	Ribeirão Pires	2003
Carapicuíba	2003	Ribeirão Preto	2005
Catanduva	2008	Rio Grande da Serra	2003
Cotia	2003	Salesópolis	2003
Diadema	2003	Santa Bárbara d'Oeste	2008
Embu das Artes	2003	Santa Isabel	2003
Embu-Guaçu	2003	Santana de Parnaíba	2003
Ferraz de Vasconcelos	2003	Santo André	2003
Francisco Morato	2003	Santos	2003
Franco da Rocha	2003	São Bernardo do Campo	2003
Guararema	2003	São Caetano do Sul	2003
Guaratinguetá	2008	São José do Rio Preto	2005
Guarulhos	2003	São José dos Campos	2005
Hortolândia	2008	São Lourenço da Serra	2003
Ilhabela	2008	São Paulo	1999/2000
Indaiatuba	2008	Sorocaba	2005
Itapeçerica da Serra	2003	Sumaré	2008
Itapevi	2003	Suzano	2003
Itaquaquecetuba	2003	Taboão da Serra	2003
Jacareí	2008	Taubaté	2008
Jandira	2003	Ubatuba	2008
Jundiaí	2008	Vargem Grande Paulista	2003
Juquitiba	2003		

Fonte: Elaboração Própria a partir de informações extraídas da SSP/SP.

Apêndice B – Descrição das categorias da CID 10 definidas como óbitos por agressões

Código	Descrição
X85	Agressão por meio de drogas, medicamentos e substâncias biológicas
X86	Agressão por meio de substâncias corrosivas
X87	Agressão por pesticidas
X88	Agressão por meio de gases e vapores
X89	Agressão por outros produtos químicos e substâncias nocivas especificadas
X90	Agressão por produtos químicos e substâncias nocivas não especificadas
X91	Agressão por enforcamento, estrangulamento ou sufocação
X92	Agressão por meio de afogamento e submersão
X93	Agressão por disparo de arma de fogo de mão
X94	Agressão por meio de disparo de espingarda, carabina ou arma de fogo de maior calibre
X95	Agressão por disparo outra arma de fogo ou de arma não especificada
X96	Agressão por meio de material explosivo
X97	Agressão por meio de fumaça, fogo e chamas
X98	Agressão por vapor de água, gases ou objetos quentes
X99	Agressão por objeto cortante ou penetrante
Y00	Agressão por meio de um objeto contundente
Y01	Agressão por meio projeção de um lugar elevado
Y02	Agressão por meio de projeção ou colocação da vítima diante de um objeto em movimento
Y03	Agressão por meio de impacto de um veículo a motor
Y04	Agressão por meio de força corporal
Y05	Agressão sexual por meio de força física
Y06	Negligência e abandono
Y07	Outras síndromes de maus tratos
Y08	Agressão por outros meios especificados
Y09	Agressão por meios não especificado

Fonte: Elaborado a partir das informações constantes no SIM/SVS/MS.

Apêndice C – Descrição das categorias da CID 10 definidas como lesões autoprovocadas intencionalmente

Código	Descrição
X60	Autointoxicação por e exposição, intencional, a analgésicos, antipiréticos e anti-reumáticos, não-opiáceos
X61	Autointoxicação por e exposição, intencional, a drogas anticonvulsivantes [antiepilépticos] sedativos, hipnóticos antiparkinsonianos e psicotrópicos não classificados em outra parte
X62	Autointoxicação por e exposição, intencional, a narcóticos e psicodislépticos [alucinógenos] não classificados em outra parte
X63	Autointoxicação por e exposição, intencional, a outras substâncias farmacológicas de ação sobre o sistema nervoso autônomo
X64	Autointoxicação por e exposição, intencional, a outras drogas, medicamentos e substâncias biológicas e às não especificadas
X65	Autointoxicação voluntária por álcool
X66	Autointoxicação intencional por solventes orgânicos, hidrocarbonetos halogenados e seus vapores
X67	Autointoxicação intencional por outros gases e vapores
X68	Autointoxicação por e exposição, intencional, a pesticidas
X69	Autointoxicação por e exposição, intencional, a outros produtos químicos e substâncias nocivas não especificadas
X70	Lesão autoprovocada intencionalmente por enforcamento, estrangulamento e sufocação
X71	Lesão autoprovocada intencionalmente por afogamento e submersão
X72	<u>Lesão autoprovocada intencionalmente por disparo de arma de fogo de mão</u>
X73	<u>Lesão autoprovocada intencionalmente por disparo de espingarda, carabina, ou arma de fogo de maior calibre</u>
X74	<u>Lesão autoprovocada intencionalmente por disparo de outra arma de fogo e de arma de fogo não especificada</u>
X75	Lesão autoprovocada intencionalmente por dispositivos explosivos
X76	Lesão autoprovocada intencionalmente pela fumaça, pelo fogo e por chamas
X77	Lesão autoprovocada intencionalmente por vapor de água, gases ou objetos quentes
X78	Lesão autoprovocada intencionalmente por objeto cortante ou penetrante
X79	Lesão autoprovocada intencionalmente por objeto contundente
X80	Lesão autoprovocada intencionalmente por precipitação de um lugar elevado
X81	Lesão autoprovocada intencionalmente por precipitação ou permanência diante de um objeto em movimento
X82	Lesão autoprovocada intencionalmente por impacto de um veículo a motor
X83	Lesão autoprovocada intencionalmente por outros meios especificados
X84	Lesão autoprovocada intencionalmente por meios não especificados

Fonte: Elaborado a partir das informações constantes no SIM/SVS/MS.

Apêndice D – Resultados do teste de autocorrelação espacial I de Moran para painel nos resíduos do modelo DID estimado por efeitos fixos.

Matriz de pesos espaciais	Coefficiente	<i>p</i> -valor
<i>k1</i>	0,024	0,242
<i>k2</i>	-0,011	0,355
<i>k3</i>	0,011	0,288
<i>k4</i>	0,035	0,031
<i>k5</i>	0,040	0,009
<i>k6</i>	0,024	0,058
<i>k7</i>	0,029	0,020
<i>k8</i>	0,028	0,016
<i>k10</i>	0,035	0,002
<i>k15</i>	0,041	0,000
<i>k20</i>	0,049	0,000
<i>distância inversa</i>	0,012	0,003

Fonte: Elaboração Própria.