Econometric Studies to the Economic and Social Factors of Crime

ISBN 978 90 361 0124 0

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **454** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

Econometric Studies to the Economic and Social Factors of Crime

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Economische Wetenschappen en Bedrijfskunde
op donderdag 7 mei 2009 om 10.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Sunčica Nikola Vujić

geboren te Belgrado, Joegoslavië

promotoren: prof.dr. S.J. Koopman
prof.dr. G.J. van den Berg

"Having looked at the structure of ordinary PhD's,
the Loneliness of the Long Distance Runner springs to mind!"
(Succeeding with your Doctorate, 2005)

Mojoj mami

Acknowledgements

"When you are in the middle of a story it isn't a story at all, but only a confusion; a dark roaring, a blindness, a wreckage of shattered glass and splintered wood; like a house in a whirlwind, or else a boat crushed by the icebergs or swept over the rapids, and all aboard powerless to stop it. It's only afterwards that it becomes a story at all. When you are telling it to yourself or someone else."
(Succeeding with your Doctorate, 2005; Citation from Attwood, 1996, p. 298).

I came to the Netherlands many rainy years ago in order to further pursue my academic career. The result of a research path that has never been very straight or easy is in front of the reader. There are many people whom I encountered on the way, who deserve to be mentioned and properly thanked.

For coming to the Netherlands in the first place, I have to thank Maarten Lindeboom and Kees van Montfort who had read my application for a Ph.D. position and decided to offer me a four-year contract. Amsterdam won the race between Newcastle, Southampton, Prague and Madrid and I packed my suitcases once again and sat on a Eurolines bus from Budapest to Amsterdam. I met Siem Jan Koopman and Gerard van den Berg on my first year in Amsterdam, whilst on the Tinbergen M.Phil. programme. They both had inspiring research ideas on the question of the relationship between crime and economics, that led me to diverge from my initial research topic and embark on a new research venture. I wish to thank Siem Jan for his eternal optimism and for being a patient, flexible, and supportive supervisor. I thank him for always finding time in his busy schedule for my questions, even complaints, and for always trying to encourage and help me. I also thank him for his friendship and personal kindness. I wish to thank Gerard for supporting my research throughout this period, for respecting my scientific integrity, and for supporting my job applications. I also thank him for offering me a shelter and a stimulating working environment at the Economics Department in the last months of putting this manuscript together. I still hope that we can finish our project on crime, unemployment, and inequality! I also thank the Tinbergen Institite and the Vrije Universiteit in Amsterdam for providing generous financial support during my Ph.D. years, for research, books, and travel to summer schools and conferences.

The fourth chapter of my thesis was completed under the guidance of Steve Machin whilst I was visiting the Centre for Economic Performance (CEP) of the London School of Economics (LSE). I thank Steve for giving me a chance to come to London and to work in the fantastic environment of the LSE on a stimulating topic of the relationship between education and crime. For funding this project, gratitude goes to the research grant of the Economic and Social Research Council. I thank my colleagues at the CEP, Olivier Marie, for helping me with the Offenders Index Database and, Jonathan Wandsworth, for tips regarding the Labour Force Survey database. At the LSE my research on economics and crime was turned from a macro to a micro focus and has remained there since this time.

The last two chapters have been done under the supervision of Dinand Webbink and Pierre Koning whilst working at the Netherlands Bureau for Economic Policy Analysis (CPB). I thank the CPB for giving me the chance to work at this prestigious Dutch research institute. I am grateful to Dinand and Pierre for their daily research support, for infinite meetings during which many ideas were born, and for teaching me how to write a research paper. I thank Petter Lundborg from the Vrije Universiteit in Amsterdam for the referee reports and his insightful comments regarding these two projects. Work at the CPB has provided me with sufficient research material to finally close the PhD chapter of my life.

Thanks to Siem Jan, Dinand, and Jacques Commandeur for helping me with the Dutch summary of this thesis and Biljana Stefanović for helping me with the Serbian summary.

During my Ph.D. years, I have been based in diverse working environments and have encountered many people whose presence in my life have indirectly, but nevertheless significantly, contributed to my years as a Ph.D. student.

On my first day at the Tinbergen Institute in Amsterdam (TIA) I met Mauro, who introduced me to my future colleagues and has made the TIA a warm and welcoming environment ever since. During those first years at the charming Tinbergen Institute on the Keizersgracht, I shared my attic room together with Mauro, Silvia, Silva, Ibolya, Jasper, Marcos, Carla, Miguel, Bert, Yiu Chung, Martijn, Boris, René, and others. Those first years in Amsterdam were marked with coursework, exams, and many lonely nights trying to meet the homework deadline for the following day. I thank Maarten Lindeboom for listening to our complaints regarding the course load and for being such a cheerful and enthusiastic Director of the Graduate Studies. During the years at the Keizersgracht, there were also many parties, lunches in the charming little cafés along the canals, and dinners at the newly opened Wagamama on the Leidseplein. People who marked those first years in Amsterdam were, next to my roommates, Maria, Aljaž, Jens, Nick, Klara, Ada, Jeroen, Joëlle, Stephan, Andrea, Ruta, Simonetta, Naomi, and Matthijs. Later also came Robert, Wendy, Sebi, Ernesto, Bea, Ana, Sandra, Gert Jan, José, Emily, Pietro, Eva, Tijmen, Razvan, Joris, and Julián. I thank Carla and Miguel for helpful Stata tips, for getting us organised for many social events, like movie nights and Museumnacht, and for making me welcome at their home in Braga. Marian, Nora, and Arianne deserve a special mention as they were always helpful and friendly secretaries.

After the Keizersgracht, I moved reluctantly to the grey surroundings of the Vrije Universiteit (VU) building close to the Amsterdam Zuid (WTC). I shared my room at the Econometrics Department with Charles, Kai Ming, Matthijs, Dinard, Joshi, Daniël, and Virginie. The VU building was ugly, but it did provide a perfect working environment. Thanks to my generous roommates, I managed to make my room at the VU look as homey as possible, with books, plants, microwave, kettle, heater, and many posters. The years that I have spent at the VU were very important for my research development and for my future career path. I thank Charles and Kai Ming for encouraging me to embark on a LaTeX path. Both Charles and Kai Ming have always found time to help me with the technicalities of my computer installations and to answer my Ox related questions. I thank Charles for generously providing a LaTeX frame to typeset this book. I thank Kai Ming for being a perfect roommate and a friend, who has provided silence or interesting conversations during those years at the VU. I thank Irma for always being helpful and friendly. I also thank Phil, who came later to the department. He showed enthusiasm for my research on economics and crime and managed to get us to the Forecasting Conference in Santander (I hope we finish that paper on risk and crime one day). Friendship with Phil, Maia, and Andreas was very important to me during those years. I also thank my other colleagues at the Econometrics Department, Gerard (van der Laan), Ben, Rein, René, Ad, Ard, Marc, Marius, Kees (van den Hoeven), Koos, Kees (van Montfort), Vitaly, Bernd, Pierre (Hoonhout), Sanne,

Angeles, Borus, and Brian.

During the years at the Amsterdam Zuidas, my special thanks goes to Joëlle and Chiara who were, for a couple of years, my VU lunch partners, and with whom I have done many other outside ventures (dinners at Joëlle's or Chiara's place, yoga with Barbara at the Zen studio, parties with Lorenzo, canoeing along the canals of Utrecht, Pinkpop festival with Martijn (van der Heide) & Co., to mention but a few). Mauro's presence, friendship, and support continued at the VU (later at the CPB as well), flavoured with many dinners at the Blasiusstraat. Special thanks goes to Mira Maletić, who has been more than just a VU project coordinator, who always helped me with formalities of the Dutch administration, and who supported me through some challenging moments during my Ph.D. Thanks to Ans and Nina, who always had a smile and time to chat along the VU corridors.

During my years as a Ph.D. student, I have spent one year at the Centre for Economic Performance at the LSE in London. Being in London and at the LSE was a dream-come-true. My heart was full of joy every morning I sat on a double-decker bus from Camden Town down to the Holborn tube station, and then that last five minutes walk along the Kingsway to the LSE. At the CEP I sat at the sector five, together with Giulia, Olivier, Teresa, Ghazala, Panu, Olmo, Joan, Sandra, Mirko (Draca), Shqiponja, Jo, Oscar, Mari, Gordon, and many others. I thank Olmo for Stata tips and parties at Chalk Farm. Mirko gave one of the nicest interpretations of my name, 'the girl who likes to sit in the sun'. Anita was constantly helpful with my computer related questions and Linda always had a cheeky joke and a big smile. Through the LSE I have also met Marieke, Moses, Lucy, and Cristina (Santos) whom I still regularly try to meet, be it in London, Amsterdam or Brussels. I thank Marieke and Moses for welcoming me to their family and for being such great people and friends (I'll always be Aldo's 'tante Zonneke'). I thank Lucy for all the cultural events in London. My London year wouldn't have been the same without the Camden Swiss Cottage Master swimming club – thanks to Lynn for teaching me butterfly, and thanks to Elizabeth, Georgia, Elena, Susan, and the rest of my swim-mates who have made London feel like home. I thank Jelena (Havelka) and Mel for making me part of their circle of friends, for fantastic parties in Canterbury and London, and for that crazy, muddy Glastonbury festival ☺!! My housemates Ludvig and Kariina deserve a special mention, for sharing with me their maisonette in the middle of Camden Town and for being such relaxed people. I have never understood complaints about the size and the buzz of London - from my Camden Town perspective, London felt just about right!

Back to the Netherlands, I had a short stop at the VU and quickly afterwards I moved to the spacious office at the CPB building in Scheveningen. My two roommates, Ben and Michiel, provided good sparring partners to practice Dutch (regarding my Dutch language improvements at the CPB, Hanneke deserves a special mention). I thank my colleagues at the CPB, Joëlle, Pierre (Koning), Dinand, Victoria (Shestalova), Viktoria (Kocsis), Victoria (Chorny-Capard), Sabine, Svetlana, Mauro, Rob, Ali, Machiel, Gijsbert, Arno, Daniël, Egbert, Nick, Ruslan, Adam, Stefan, Hugo and others for introducing me to the to the Dutch culture of a 'lunch at 12:30h', lunches at the ZO, for many nice 'uitjes', birthday cakes, yoga at the basement of the CPB building, and short after-lunch walks ('ommetje') around the 'Scheveningse bos'. The CPB building wasn't really around-the-corner from my place in Amsterdam, but I loved my spacious room full of light, with the smell of sea in the air, and the sound of seagulls from the outside.

The last months of setting this manuscript together I have spent at the Economics department at the VU, where I shared an office with Hans Bloemen. Hans was a perfect, quiet roommate during those last, arduous months. I also thank Pieter, Petter, Bas, Maarten, Mauro, Jonneke, Marta, Sumedha, Ronald, Tibor, and my other colleagues at the Economics Department for providing a motivating and socially pleasant working environment. It's only here that I saw the light at the end of a tunnel called ‘the path towards a Ph.D.’ I treasure those last months at the VU, I wish that my research career has always been so smooth, focussed, and effective.

I have encountered many other people while living in Amsterdam. Special thanks goes to my ex-Yugoslav crowd: Marijana and Peca, Žika and Tijana, Zoran and Tanja, Mirko and Olivera, Brane and Sandra, Andrej and Aleksandra, Dejan and Danijela, Ana and Peđa, Aljaž, Ana (Miškovska-Kajevska), Desa, Bora, Natalija, Mija, Ivana, Dragana, Srđan, Saša, Nikoleta, Nebojša, Branka, Vesna, Goran, Aleksej, Richard, Helena, Tanja (Đaković), and many others (like Pim and Stefan, with strong Balkan affinities). I thank them for all the parties, barbecues, Balkan film festivals, helping me to move in Amsterdam so many times, all the tips about where to buy cabbage leaves for ‘sarma’ or ‘vegeta’ (now also in Albert Heijn), and discussions about the ‘verblijfsvergunningen’ and other peculiarities of a life of a person from the Balkans abroad. I specially thank Marijana and Peca for their friendship and the warm environment of their home(s) in Diemen and de Pijp.

Another person who coloured my life in Amsterdam was my flatmate Maciek, during that one year of living in de Pijp – thanks for being so tolerant and for helping me organise the best birthday party ever! Thanks to Sigrid, Vera, Shita, and Albertine, my classical ballet mates. Thanks to Annique, Ramach, Maya, Sisi, Sara, Marije, and Tali for dinner parties, movies, and the feeling of being welcome in Amsterdam. Thanks to James for forwarding me the add for a Research Assistant at the LSE. Thanks to Viktoria for embracing me into her and Zsolt’s circle of friends – I always loved meeting up with Gitte, Jair, Bix, Patxi, and Amandine. Thanks to the ‘Italian mob’ – it has changed over the years, but it has always been there: Mauro, Silvia, Chiara, Lorenzo, Olmo (in London), Cristina (Pacciarella), Giovanni, and Francesco. Thanks to the pilates studio Jelena Petrović (special thanks to Jelena and Anna) and my Master swimming club Upstream.

During the course of my Ph.D. years I have also met many nice people during conferences and summer schools in Amsterdam and abroad. Thanks to Rafael, one of the few criminologists I have met on my way. I still have a great story to tell after our guided tour through the Red Light District of Amsterdam. Thanks to George Saridakis for inviting me to come to the conference in Tübingen. Thanks to Shushanik, Pablo and Cédric, we have made a natural team during that Copenhagen summer school on cointegration. Thanks to the people whom I met on those two summer schools in Steyr: David, Jesper, Cecilia, Salvatore and my ‘girl team’ Irina (thanks for all emails of support), Isabel, Marija, and Ralitsa. Thanks to Krisztina and Istvan for their friendship and a warm and harmonious home in Freiburg – I’ll also try one day to sort my books in the colour of the rainbow. Thanks to Arnisa and Taulant, for their friendship and kind words. Thanks to Gita and Estere, for all surprise visits and ‘good-old-times’ in Budapest. Thanks to Kati, for being a great friend, for making me part of her Hungarian circle of friends (Renata & Co.), for welcoming me to her family (thanks to Agnes, Andras, and Lazi), and for always having the door of her home in Budapest open for me.

I would have never ended up on a serious research path in the field of economics if it wasn’t for my Professors at the Economics Faculty of the University of Belgrade, who inspired me with their personalities, broad interests, and whose lectures made economics sound challenging. They have built my economics and econometrics foundation, which proved to be a valuable piece of luggage all these years abroad. I thank Milić Milovanović, Dana Popović, Milena Jovičić, Branislav Boričić, Ljubomir Madžar, Zorica Mladenović, Mihail Arandarenko, and Mlađan Dinkić, among others. They deserve all my respect and admiration. Here I also wish to mention Marina Mladenović-Komatina, who was the head of my first research department at the Serbian National Bank after I had graduated from the University of Belgrade, and who excitedly talked to me about the Kalman filter and the interesting research done at the Dutch Central Bank in Amsterdam.

Belgrade and my friends there have a very special place in my heart. I have never stopped missing just about everyone in Belgrade while living abroad. My regular visits home have put my Ph.D. thesis in a more realistic perspective and I always knew that people in Belgrade would value me for who I am, whether I succeeded on the Ph.D. path or not. Here I wish to thank Srđana and Nebojša (Srđo, you are the 'milkman of human kindness'), Biljana (Stefanović) (model student I've always wanted to be), Ivana, Tanja (Paunović), Desa, Boban, Maksa, Maja, Biljana (Soldatović), Mirjana, Dejan and Luka, Snežana and Zoki, Ivančica and Zoran, Sofija, Mirče, mala Sunčica, Slavica, Kaća and Marko ('najbolji prijatelji'), Suzana and Goran, Snežana and Vule, Jelena (Tasić), Jelena (Havelka), Snežana (Marković), Biljana (Jelić), Bojana, Natalija, Aleksandra (Borović), Igor, Tanja (Parac-Vogt), Sonja, Jelena (Aleksić), Miša, Sveta, Aleksandra (Dogramadži), moja sestra Slađa, tetka Savka, and many others. I thank you for being such wonderful people. I also thank doctor Smiljević and Živkica – I couldn't have endured all the long hours of sitting behind a computer if it wasn't for your wonderful exercises.

I thank my father, for having a keen eye for everything that goes on around him. I owe my broad interests and adventurous spirit to him. He would have been proud of me today, as he has always been. I thank my mother for her unreserved love and support, and for all those jars of 'ajvar' and jam, that kept me going through the years. I thank her for always giving me the freedom to go my own way and for always trusting my judgement. I dedicate this book to her.

At the end, I wish to thank Amsterdam, for being such a wonderful city to live in. My life would have been much more miserable hadn't it been for concerts in Paradiso and Melkweg, movies at the Kriterion, Rialto, Cavia, Uitkijk, Filmmuseum, Cinecenter and the Movies, cafés where I have always felt at home (Villa Zeezicht, Vertigo, De Jaren, De Balie, Singel 404, Wijnand Fockink, Gary's Deli, Foodism, to mention but a few), endless festivals and outdoor events, and alternative places like OT301, OCCII, Nieuwe Anita, ADM Terrain, and Ship of Fools. Amsterdam does not have the grandiosity of cities like London, Budapest or Paris, but it has charming canals, that turn yellow at night, along which I have been biking all these years.

In Amsterdam, 11th of March, 2009

Contents

Acknowledgements **vii**

Contents **xiii**

1 Introduction **1**
1.1 Motivation and background . . . 1
1.2 Outline . . . 2
1.2.1 The impact of sentence reform in Virginia . . . 2
1.2.2 Economic trends and cycles in crime . . . 4
1.2.3 Education and crime . . . 5
1.2.4 Early behavioural problems and human capital . . . 7
1.3 Software . . . 8

2 Intervention Time Series Analysis of Crime Rates: The Impact of Sentence Reform in Virginia **9**
2.1 Introduction . . . 9
2.2 Changes in criminal justice system of Virginia . . . 11
2.3 Plan and details of empirical intervention study . . . 17
2.3.1 Data description . . . 17
2.3.2 Intervention analysis . . . 20
2.3.3 Preliminary data analysis . . . 22
2.4 Preliminary approaches to intervention analysis . . . 22
2.4.1 Regression analysis . . . 22
2.4.2 Regression models with ARMA errors . . . 27
2.4.3 Box-Jenkins ARIMA model approach . . . 30
2.5 The structural time series approach . . . 34
2.5.1 The basic structural time series model . . . 34
2.5.2 State space representation, filtering and smoothing . . . 37
2.5.3 Equivalence between STS and ARIMA models . . . 42
2.5.4 Structural time series analysis results . . . 43
2.6 The multivariate structural time series approach . . . 51
2.6.1 The multivariate STS approach: state space formulation . . . 51
2.6.2 Multivariate STS intervention analysis results . . . 55
2.7 General discussion of empirical findings . . . 56
2.7.1 Relation to difference-in-difference approach . . . 56

2.7.2 Review of empirical findings 61
2.A Appendix 67
2.A.1 Diagnostic and Goodness-of-Fit Tests 67
2.A.2 Definitions of the analysed crime categories 69

3 Economic Trends and Cycles in Crime: A Study for England and Wales 71
3.1 Introduction 71
3.2 Decomposition of the economic and crime series 75
3.2.1 Data and methodology 75
3.2.2 Trend-cycle decompositions for economic and crime series 77
3.2.3 Regression effects 78
3.3 Multivariate trend-cycle decompositions 83
3.3.1 The multivariate unobserved components model 83
3.3.2 Multivariate decomposition results 84
3.3.3 Multivariate decomposition results with explanatory variables . . . 87
3.3.4 Model for thefts 87
3.3.5 Multivariate model with economics, crime and police 90
3.4 Literature overview on crime and economic incentives 94
3.4.1 Criminal opportunity and criminal motivation 94
3.4.2 Unemployment and crime 96
3.4.3 Income and crime 99
3.5 Conclusion 101
3.A Appendix 102
3.A.1 Definitions of variables and sources of data 102

4 Crime and Education: New Evidence From Britain 103
4.1 Introduction 103
4.2 How education can impact on crime 104
4.3 Crime data sources and schooling reform in Britain 107
4.3.1 Crime data sources 107
4.3.2 Background on the schooling reform in Britain 108
4.4 Data description and matching approaches 109
4.4.1 Data 109
4.4.2 Approaches 109
4.5 Empirical approach 112
4.5.1 The statistical model 112
4.5.2 Logit models from Census data 115
4.6 Estimation results 115
4.6.1 Approach 1 115
4.6.2 Approach 2 123
4.7 Conclusions 129
4.A Appendix 131

5 Why Are Criminals Less Educated than Non-Criminals? Evidence from a Cohort of Young Australian Twins 133
5.1 Introduction . . . 133
5.2 Previous studies . . . 135
5.3 Empirical strategy . . . 136
5.4 Description of data . . . 138
5.5 The effect of early arrests on educational attainment . . . 142
5.6 The effect of human capital on crime . . . 146
5.7 Robustness . . . 149
5.7.1 Missing values due to the routing of the questionnaire . . . 149
5.7.2 Measurement error . . . 149
5.8 Decomposing crime–education relationship . . . 153
5.9 Conclusions and discussion . . . 154
5.A Appendix . . . 156
5.A.1 Variable definitions . . . 156
5.A.2 Robustness checks tables . . . 156

6 The Effect of Childhood Conduct Disorder on Human Capital: Evidence from a Cohort of Young Australian Twins 159
6.1 Introduction . . . 159
6.2 Previous studies . . . 161
6.3 Data . . . 163
6.3.1 Conduct disorder . . . 164
6.3.2 Positive and negative human capital . . . 165
6.3.3 Descriptive statistics . . . 166
6.4 Methodology . . . 170
6.5 Main estimation results . . . 172
6.5.1 The effect of conduct disorder on human capital . . . 173
6.5.2 The effect of the timing of conduct disorder behaviours . . . 181
6.5.3 Summary of the timing of conduct disorder . . . 184
6.6 Robustness checks . . . 184
6.6.1 Missing values due to the routing of the questionnaire . . . 184
6.6.2 Excluding pairs of twins with large differences in education . . . 185
6.6.3 Measurement error in conduct disorder . . . 187
6.6.4 Spill-over effects within pairs of twins . . . 187
6.6.5 Summary of robustness checks . . . 189
6.7 Conclusions . . . 189
6.A Appendix . . . 192
6.A.1 Variable definitions . . . 192
6.A.2 APA DSM-IV diagnostic criteria for conduct disorder . . . 193

7 Conclusions 195

Bibliography 199

Samenvatting (Summary in Dutch) 213

Sažetak (Summary in Serbian) **221**

List of Tables

List of Tables **xvii**

2.1 Comparison of median time-served (in years) in 1993 (system with parole) and anticipated median time-served for offenders sentenced in 2001 (system without parole) . . . 14
2.2 Major Criminal Justice Initiatives in Virginia, 1988-1998 . . . 15
2.3 Estimated interventions for regression models . . . 28
2.4 Estimated interventions for a regression plus AR error models . . . 31
2.5 Estimated interventions for ARIMA models . . . 33
2.6 Special Cases of the STS Models . . . 35
2.7 Fully Specified LL Model . . . 37
2.8 Estimated interventions for structural time series models - Property crimes 49
2.9 Estimated interventions for structural time series models - Violent crimes . 50
2.10 Estimated interventions from structural time series models for both property and violent crimes . . . 52
2.11 Estimated interventions for multivariate STS models - Burglary as a treatment series . . . 57
2.12 Estimated interventions for multivariate STS models - Robbery as a treatment series . . . 58
2.13 Estimated interventions for multivariate STS models - Murder as a treatment series . . . 59
2.14 Estimated interventions for multivariate STS models - Rape as a treatment series . . . 60
2.15 Crime rates in Virginia and the rest of the United States, 1991-2003 . . . 62
2.16 Estimates of the impact of the parole abolition and sentence reform on Virginia crime rates . . . 63

3.1 Maximum likelihood estimation of univariate models for real GDP, Unemployment, Burglary and Theft (sample 1955-2001, without and with regression effects) . . . 79
3.2 Maximum likelihood estimation results for multivariate decomposition model with real GDP, Unemployment and Burglary (sample 1955-2001, without and with regression effects) . . . 88

3.3 Maximum likelihood estimation results for multivariate decomposition model with real GDP, Unemployment and Theft (sample 1955-2001, without and with regression effects) . . . 91
3.4 Maximum likelihood estimation results for multivariate decomposition model with real GDP, Unemployment, Burglary, Theft, and Police (sample 1955-2001, without and with regression effects) . . . 92
3.5 Maximum likelihood estimation results for multivariate decomposition model with real GDP, Unemployment, Burglary, Theft, and Police (sample 1955-2001, without and with regression effects, continued) . . . 93
3.6 Cantor and Land (1985): Motivational and opportunity crime effects . . . 95

4.1 Summary statistics, Approach 1 . . . 111
4.2 Property crime convictions and years of education . . . 117
4.3 Property crime convictions and no educational qualifications . . . 119
4.4 Property crime convictions and no educational qualifications, males separately . . . 121
4.5 Property crime convictions and no educational qualifications, females separately . . . 122
4.6 Violent crime convictions and no educational qualifications, males separately 124
4.7 Checks of robustness . . . 125
4.8 Imprisonment rates (percent), 2001 Census . . . 127
4.9 Logit estimates of imprisonment equations . . . 128
4A.1 Type of communal establishment, England, Wales and Scotland, Census 2001 . . . 131

5.1 Summary statistics for the main estimation samples . . . 141
5.2 Arrests by schooling level(%) . . . 142
5.3 Estimates of the effect of early arrests on educational attainment . . . 144
5.4 Estimates of the effect of the timing of the early arrest on educational attainment . . . 145
5.5 Estimates of the effect of high school completion on crime . . . 148
5.6 Estimates of the effect of early arrests on educational attainment after imputations for missing values on early arrests . . . 150
5.7 Estimates of the effect of senior high school completion on crime after imputing missing values on crime measures . . . 151
5.8 IV estimates of the effect of senior high school completion on crime . . . 153
5.9 Fixed effect estimates of the effect of education on the probability of being arrested . . . 154
5A.1 Estimates of the effect of years of education completed on crime . . . 157
5A.2 IV estimates of the effect of years of education on crime . . . 158
5A.3 IV estimates of the effect of senior high school completion on crime after imputing missing values on crime measures . . . 158

6.1 Means table for sample of children with all conduct disorder measures non-missing . . . 167
6.2 Distribution of conduct disorder score (% of twins with each score) . . . 168

6.3 Means of conduct disorder categories 169
6.4 Estimates of the effect of conduct disorder on positive human capital . . . 174
6.5 Estimates of the effect of conduct disorder on negative human capital . . . 175
6.6 Estimates of the effect of more than 5 conduct disorder behaviours (above 90th percentile) 177
6.7 Estimates of the effect of at least 10 conduct disorder behaviours (above 98th percentile) 178
6.8 The effect of four subscales of conduct disorder on positive human capital . 179
6.9 The effect of four subscales of conduct disorder on negative human capital 180
6.10 The effect of the timing of conduct disorder on the conduct disorder score . 181
6.11 The effect of the timing of conduct disorder on positive human capital . . . 182
6.12 The effect of the timing of conduct disorder on negative human capital . . 183
6.13 Estimates of the effect of conduct disorder (APA definition) on positive and negative human capital after imputation of missing values due to the routing of the questionnaire 185
6.14 Estimates of the effect of conduct disorder (APA Definition) on positive and negative human capital after excluding pairs with extreme education differences 186
6.15 IV-estimates of the effect of conduct disorder on positive and negative human capital 188
6.16 If one twin has a problem, does the other twin have it also? (Percent) . . . 189
6.17 Random effect estimates of own and sibling's conduct disorder on human capital 190

List of Figures

List of Figures **xxi**

2.1 Time served in prison under parole system and truth-in-sentencing system 16

2.2 Property crime rates: burglary, larceny, motor vehicle theft, and robbery. . 18

2.3 Violent crimes rates: aggravated assault, murder, and rape. 19

2.4 Intervention effects: (a) pulse intervention, modelled as an additive outlier; (b) step intervention, modelled as a level break; (c) smooth step intervention, modelled as a gradual level break. 21

2.5 Mean change in property crime rates: (a) burglary; (b) larceny; (c) motor vehicle theft; (d) robbery. 23

2.6 Mean change in violent crime rates: (a) aggravated assault; (b) murder; (c) rape. 23

2.7 Trend change in property crime rates: (a) burglary; (b) larceny; (c) motor vehicle theft; (d) robbery. 24

2.8 Trend change in violent crime rates: (a) aggravated assault; (b) murder; (c) rape. 24

2.9 Trend change in property crime rates. Subsample 1991:01-1998:12: (a) burglary; (b) larceny; (c) motor vehicle theft; (d) robbery. 25

2.10 Trend change in violent crime rates. Subsample 1991:01-1998:12: (a) aggravated assault; (b) murder; (c) rape. 25

2.11 Trend and seasonal components in property crime rates: (a) burglary; (b) larceny; (c) motor vehicle theft; (d) robbery. 45

2.12 Trend and seasonal components in violent crime rates: (a) aggravated assault; (b) murder; (c) rape. 46

2.13 Auxiliary residual plots of the LL and LLT fitted models: Property crimes 47

2.14 Auxiliary residual plots of the LL and LLT fitted models: Violent crimes . 48

3.1 Yearly time series for England and Wales between 1955 and 2001, in logs: (i) real gross domestic product (GDP); (ii) number of unemployment per capita; (iii) number of reported burglaries per capita; (iv) number of reported thefts per capita. 76

3.2 Trend-cycle decompositions are based on the univariate unobserved components model (3.1), (3.2) and (3.4). The estimated decompositions are presented for (row-wise) real GDP, Unemployment, Burglary and Theft. The estimated components are for (column-wise) trend, short cycle and medium cycle. The estimation is carried our for the sample 1955-2001. . . 80
3.3 Estimated cycles (scale-adjusted by range) are based on the univariate unobserved components model (3.1), (3.2) and (3.4). The first plot presents the short cycles for real GDP, Unemployment, Burglary and Theft. The second plot present the medium cycles for the same series. The estimation is carried our for the sample 1955-2001. 80
3.4 Estimated multivariate decomposition for time series – row-wise – B (burglary), U (unemployment) and G (real GDP) with components – column-wise – (i) trend (with observations), (ii) short cycle, (iii) medium cycle and (iv) irregular. 89

4.1 Percentage left full-time education aged 14 and 15 by year of birth 114
4.2 Property conviction rates (per 1000 population) by age, 1984 to 2002 . . . 116

5.1 Age of first arrest by schooling level 143

6.1 The associations between conduct disorder scores and human capital . . . 170

Chapter 1

Introduction

"Today, we take a break from trade balances, oil prices, interest rates and the like to consider studies with surprising social and cultural content. They focus on specific parts of life where rationality is usually not an issue but where, in the shadows, homo economicus may be at work. In other words Wackonomics is an extreme form of 'Economic Imperialism'. Economic imperialism means that one applies the tools of economics on subjects or questions that aren't traditionally considered as belonging to economics."
– Wackonomics, The Economist, 28 October 2000

1.1 Motivation and background

From one election to another, crime remains one of the biggest social and political issues, and it is a favourite subject of campaign promises by politicians. Crime imposes large costs onto society which are not taken into account by official measures of well-being and happiness of the nations. In addition, crime responds to economic conditions and incentives, which is well accepted among economists ever since the seminal papers by Garry Becker (Becker, 1968) and Isaac Ehrlich (Ehrlich, 1973). According to the standard economic model of crime, an individual's criminal activity depends on total income from legal activities, the preferences of the individual, and exogenous factors that affect the probability of arrest and the sanctions imposed. Despite the substantial work in the area, the public seems largely unaware of the economics view of crime and its implications.

From an economic perspective, crime is easily described in a supply-and-demand framework in which criminals supply crime, the public demands protection from crime, and the government provides public protection (Becsi, 1999). The supply of crime is seen as a choice between legitimate activities and work on the one hand and criminal activities on the other. The choice depends on the net benefit to crime, which is the benefit of the criminal activity above all other costs associated with the crime. Eide (2000) gives a thorough overview of benefits and costs of crime. Gaines obtained from criminal activity depend on the type of crime and the individual criminal, and can be monetary or psychic (thrill of danger, peer approval, sense of accomplishment or satisfaction of wants in case of rape). Costs of crime include direct and indirect (opportunity) costs. Direct costs of crime are material costs (equipment, guns, vehicles), psychic costs (guilt, anxiety, fear, dislike or risk), and expected punishment costs (fines, incarceration, and other sanctions, as well as pecuniary costs arising from lawsuits). Opportunity costs of crime consist of net benefit (benefit minus cost) of the legal activity forgone while planning, performing, and concealing the criminal act. For example, education and welfare might be thought to increase the opportunity cost of committing crime by increasing legitimate earnings. The supply of crime is positively related to the net benefit to criminal activities, meaning that criminals will increase their activities when the net benefit rises. Examples of conditions that might cause a rightward shift of the supply curve include demographics (a higher proportion of youth), fewer employment opportunities at a given wage, and reductions in imprisonment. While property crimes might be thought of as most responsive to economic conditions, many violent crimes are committed as by-products to crimes for economic gain and thus are also explainable with economics.

This dissertation mainly focuses on the supply side of crime, analysing the relationship between crime and economic incentives using aggregate (national) time series data (US and UK), aggregate cross-sectional data (UK), and individual survey data (Australia). The main purpose of this study is to show empirical evidence that if we consider crime as a rational behaviour, then economic and social levers should be the main tools of a society in order to prevent and combat crime.

1.2 Outline

1.2.1 The impact of sentence reform in Virginia

Chapter 2 focuses on crime trends in the United States, particularly in Virginia. Historically, crime trends in the US have been increasing through the 1960s and 1970s, have radically been reversing the trend through the 1980s, in order to see a dramatic fall in crime in almost all categories in the 1990s. Levitt and Dubner (2005) writes that the 1960s and the 1970s were “the heyday of a liberal justice system and the criminal’s right movement” – the likelihood of punishment was so low that it simply did not cost very much to become a criminal. However, by the 1980s this trend has radically reversed. Criminals’ rights were limited and stricter sentencing guidelines were put in place. During the 1990s, almost all states in the US adopted one or more mandatory sentencing laws. They all had one thing in common: they all imposed mandatory, statutory increases in prison sentences on individuals who were already going to be imprisoned. However, there is still a large debate among criminologists and policy makers as to which factors contributed to the crime-drop in the US in the 1990s (e.g., legalized abortion 20 years earlier, expanding economy, community policing, changes in crack and other drug markets, and/or higher arrest and incarceration rates). Levitt and Dubner (2005) write that a sharp drop in crime throughout the US in the 1990s have been mainly explained by increased reliance on prisons, increased number of police officers per capita, the crash if the crack market, and the legalisation of abortion. This chapter tries to contribute to better understanding of the statistical relationship between anti-crime efforts and crime reductions over time, using data for Virginia, in the period from 1984 to 2006.

We examine the impact of the new legislation in Virginia on reported property (burglary, larceny, motor vehicle theft, and robbery) and violent (aggravated assault, murder, and forcible rape) crime rates using different time series approaches. Virginia abolished parole and reformed sentencing for all felony offenders committed on or after January 1, 1995. The net result of the implementation of the new legislation was a substantial increase in the sentences for the violent offences (especially rape and murder) and also for offenders with a violent past. Although property crimes have not been directly targeted by this new legislation, increases in prison sentences have also been anticipated for some categories of property crimes (burglary and robbery in particular). To the extent that severity of punishment serves as a deterrent to committing crimes, we would expect that the Virginia's anti-crime initiatives reduced the reported crime rates, especially for violent offences.

Virginia's abolition of parole and reform of the sentencing system provides a useful social experiment to study. First the legislation was very sweeping and impacted all felonies. Further, such sweeping legislation was enacted at a time in which there were very large (and favourable) changes in a number of social and economic indicators. Finally, the 1990s also saw the implementation of a number of initiatives focused on reducing crime at the Federal, State and Community levels. Disentangling the impact of parole abolition from the other factors poses multiple design and analytical challenges.

We examine the impact of this new legislation in Virginia on reported crime rates using different time series approaches. In particular, structural time series models and Box-Jenkins ARIMA models are considered. Both time series approaches provide effective methodologies for intervention analysis. The new legislation is modelled as the level shift or step intervention, where the value of the level of the crime time series instantly changes at the time point when the intervention takes place, and where the level change is permanent after the intervention. The empirical literature provides some but limited support for the deterrent impact of parole abolition and sentence reform in Virginia. Further, Kessler and Levitt (1999) report that previous research on the impact of sentence enhancements on reported crime rates in the United States has reported conflicting findings and often find no effect (see for example Loftin et al., 1983). We confirm these findings based on univariate time series models. We also consider a parsimonious framework of multivariate structural time series models, which allow the inclusion of control groups in the analysis. We illustrate the flexibility of the multivariate approach and provide new evidence for the deterrent impact of the new legislation.

This chapter contributes to the existing literature as follows. First, we see it as a comprehensive evaluation of the new legislation. According to the Virginia crime officials, research to date was unable to determine if the observed reductions in crime rates were due to specific anti-crime initiatives. Second, we view our work as an introduction of structural time series (STS) methods to crime data analysis. To our knowledge, the STS estimation approach has not been used extensively in crime analysis. On the other hand, the regression and especially the ARIMA models have been widely used in the criminal justice literature. Hence, we view the present chapter as a potential contribution to time series methodology in criminology. Third, when dealing with time series, proposed models for analysing the intervention effects of parole abolition and sentence reform in Virginia clearly favour ARIMA and/or structural time series approaches to modelling intervention.

1.2.2 Economic trends and cycles in crime

Chapter 3 focuses on the crime trends in the United Kingdom. Particularly, we look at the effect of a business cycle on property crime rates (burglary and theft) in England and Wales throughout the second half of the twentieth century. In the UK crime also increased rapidly through the 1970s and 1980s, reaching record levels by the start of the 1990s, becoming an important public policy issue. The increase was far greater than that experienced in the rest of Europe. Moreover, property crime rose rapidly in the UK through the 1970s, 1980s, and early 1990s. Literature shows that the level of property crime in England and Wales was higher than in the US through the mid-1990s (Machin and Meghir, 2004).

The business cycle has a pervasive effect on the structure of economic opportunity and hence on behaviour. The effect is reflected in social indicators as diverse as school enrolments, birth rates, and labour force participation. It would be then surprising if crime rates were immune to general business conditions. Cook and Zarkin (1985) write that "the relationship between general business cycle condition and crime has been a popular subject of study for empirical criminologists since the mid-nineteenth century," particularly in England and Wales (see for example Swaine-Thomas, 1927 and Wolpin, 1978). Literature suggests that property crimes have a counter-cyclical pattern, falling and rising along with the ups and downs of the cycle in economic activity, while violent crimes are not so much business cycle driven (except maybe for robbery, which has a clear for-profit motivation).

In order to study economic-induced behaviour in recorded burglary and theft data in England and Wales in the period from 1955 to 2001, we adopt both univariate and multivariate time series frameworks within the unobserved components models. We explicitly model the business cycle as a stationary autoregressive moving average (ARMA) process, which is a novelty of our approach as compared to the existing examinations of the effect of a business cycle in the crime economics literature. In the multivariate structural time series framework, we have simultaneously modelled the time series of burglary and theft together with the time series of real GDP, unemployment, and police. The second contribution of this chapter is the simultaneous analysis of crime (burglary and theft), economic (GDP and unemployment), and deterrence (police) time series. We also account for various other variables that can explain variations in crime rates, like criminal justice variables (conviction rate, sentence length, and imprisonment), demographic variables (male, youths aged 15-24 years), the effect of a change in recording practice (a dummy variable accounting for the introduction of the Theft Act in 1968), and the effect of number of police officers. We find strong evidence for cyclical behaviour in burglary and theft crime rates. However, we find no evidence of a separate crime cycle, it can be fully explained by economic cyclical behaviour. The explanatory variables affect more the short-term dynamics than the long-term dynamics.

A seminal paper by Cantor and Land (1985) built foundations for the macroeconomic link between economic conditions and crime. They propose two paths through which the business cycle could affect the level of crime. Criminal motivation works through the impact of changing economic conditions on social strain and social control, whereas criminal opportunity affects the availability and vulnerability of criminal targets through routine activities theory. The two effects run in opposition to each other and they need not occur at the same time. The third contribution of this chapter is that the model detects the motivational and opportunity effects of burglary and theft in relation to economic conditions.

1.2.3 Education and crime

Application to England and Wales

Crime is widespread among young people, particularly young men. Examples in the literature abound. Rutter et al. (1998) report that a significant minority of young people will go on to acquire a criminal record at some point in their adolescence. The UK Home Office analysis of two cohorts of people born in 1953 and 1958 shows that approximately a third of male adults had been convicted of at least one standard list offence by their thirties (Home Office, 1995). Many of these convictions were as a result of offences committed before adulthood. In their longitudinal study of 411 males in inner London followed from age 7 into adulthood, Farrington et al. (1998) report that the average criminal career began at some point between 14 and 21 years, lasted ten years, ended at about age 26, and (for those who committed more than one offence) consisted of 4.6 offences leading to conviction. This relationship between age and crime is robust over time, across countries, demographic subgroups, types of crime, and holds irrespective of the way the crime is measured (Hirschi et al., 1983). In the US these numbers are similar. Grogger (1998) writes that thirty-five percent of all Philadelphia males born in 1945 were arrested before the age of 18, and one-third of all California men born in 1956 were arrested between the ages of 18 and 30. The 1990 US Census counted 1.1 million persons in jail, the vast majority of who were men in their twenties and thirties.

In chapter 4 we argue that one of the main causes of high numbers of crime among young people are due to the early school drop-out and reduced years of education, which then have repercussions on individual's labour market productivity and legal versus illegal income opportunities. Conceptually, there are difficulties in isolating the impact of education. Specifically, it is difficult to guarantee that the direction of causation flows from education to crime (and not the other way round). To address this problem we adopt a quasi-experimental approach relying on variations in education induced by changes in compulsory school leaving age laws in England and Wales over time to validate the direction of causation. We look at the relationship between crime and education using two British data sources. The first, the Offenders Index Database (OID) covers all convictions in England and Wales, and we match this to Labour Force Survey data on education for age cohorts over time. The second is data on imprisonment from the 2001 Census, where we look at crime and education in a large cross-section of the British population.

Our results show sizeable effects of education on crime. Moreover, empirical estimates from the instrumental variables strategy that we adopt are, when we use an education variable that is best suited to this approach, rather similar to those that are just based on ordinary least squares regressions that may not have a causal interpretation. In our empirical models of property crime convictions, we report that having low education levels, especially possessing no educational qualifications, is significantly associated with higher levels of offending. We corroborate this with cross-sectional findings on imprisonment and lack of educational qualifications from the 2001 Census. The implications of these findings are clear and they show that improvements in human capital accumulation through the education system or other means that can be showed to be effective for enhancing individual labour market productivity and should be thought of as a key policy tool in the drive to reduce crime. In words of Eide (2000), "young people are perhaps not different, just poorer."

Application to Australia

In chapter 5, we address the relationship between human capital and crime using data of a sample of young Australian twins. We exploit two aspects of the Australian survey data on education and crime. First, because the data are obtained from twins, we are able to control for many unobserved characteristics affecting both criminal behaviour and schooling decisions. Second, because criminal behaviour is measured over different periods of time – prior to and after senior high school completion – we can address the causality between crime and education as well. Since early criminal behaviour may affect human capital formation, and human capital may influence criminal behaviour in later stages of life, we follow a two step analysis.

First, we study the relationship between early crime and the accumulation of human capital. In particular, we estimate the effect of arrests before the age of 18 on educational attainment by using within-twin estimation. In addition, we investigate whether the timing of the arrest matters for educational attainment. Second, we estimate the effect of educational attainment on three measures of crime: incarceration, arrests since the age of 18, and number of arrests. As early criminal behaviour might be an important confounder in the estimation, we control for early arrests and measures of conduct disorder within pairs of twins.

This chapter contributes to the economic literature on the relationship between education and crime in several aspects. First, we add to the not so abundant literature on the causal relationship between education and crime and use an identification strategy that has not been applied in this context before – that is, we exploit the longitudinal nature of our data in order to estimate the relationship between human capital and crime in both directions. Second, we investigate the effect of early criminal behaviour on investment in human capital while controlling for fixed effects within pairs of twins. We are not aware of studies in the economic literature that estimate the causal effect of early criminal activity on educational attainment. Third, there is growing interest in the economic literature over the effects of early conditions in life on adult outcomes (Currie and Stabile 2006, 2007; Borghans et al., 2008). This chapter addresses similar issues.

We find that early arrests (arrests before the age of 18) have a strong effect on human capital accumulation. These effects are largely driven by the timing of the early arrest, showing that arrests at age 13, 14 or 15 are most detrimental for educational attainment. We also find human capital to reduce crime. However, when controlling for early arrests and early behaviour problems, the estimated effect of human capital on crime reduces to less than a quarter of the previously estimated association. The strong detrimental effects of early criminal behaviour become also transparent if we consider the estimated effects of early arrests on all three measures of crime. We then find large effects of early criminal behaviour on participation in crime later on. These effects are much larger than the (isolated) impact of human capital on crime. We conclude that the strong association between human capital and crime is mainly driven by the effect of early criminal behaviour on educational attainment. Programs that succeed in preventing early criminal behaviour might yield high social and private returns.

1.2.4 Early behavioural problems and human capital

Benefiting further from the data on Australian twins, in chapter 6 we analyse the effect of early behavioural problems on human capital accumulation, where we see crime as a negative human capital ("building capital in anti-social or criminal activities," Currie and Stabile, 2007). Literature shows that many children have mental health problems which hinder their normal development and functioning. Anxiety, conduct disorder, attention deficit hyperactivity disorder, and depressive disorders are the most common. For example, up to 500,000 (14%) of Australian children between the ages of 4 and 18 years have significant mental health problems (Australian Centre for Community Child Health, 2006); up to 50% of 'hard to manage' pre-school children are likely to have continuing hardships requiring professional help; approximately one in five children and adolescents in the US may have a mental health disorder (Currie and Stabile, 2006). Despite these large numbers little is known about the longer term effects of these mental health problems of children.

This chapter analyses the longer term effects of childhood conduct disorder on human capital accumulation and violent and criminal behaviour later in life using data of Australian twins. If someone shows "a repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate societal norms or rules are violated," then he/she falls into a category of individuals with a conduct disorder problem. Conduct disorder is known as a 'disruptive behaviour disorder' because of its impact on children and their families, neighbours, and schools, and is largely associated with delinquent or criminal activity. We measure conduct disorder using diagnostic criteria from the American Psychiatric Association (APA, 1994). In line with Currie and Stabile (2007) we estimate effects on positive human capital, including measures such as grade repetition, marks in primary and high school, and educational attainment, and on 'negative' human capital, including measures such as being arrested, spent time in jail, and physically attacking others.

The two main challenges in estimating the longer term effects of childhood conduct disorder on human capital are the measurement of conduct disorder and omitted variable bias. This chapter contributes to the economic literature by addressing these two main challenges. First, it is easy to know when a child has a fever but a child's mental health problem is harder to identify. We use self-reports of adult twins on 21 statements that follow the definition of conduct disorder according to the American Psychiatric Association criteria. Another contribution of this chapter is that we also have information on the age at which the conduct disorder behaviours occurred. We use this information for investigating the effects of the timing of the problem behaviours on human capital. Second, estimates of the effect of conduct disorder on human capital might be confounded by unobserved differences between children and their families. In this chapter we estimate within-family models using data on (identical) twins. The advantage of using twins instead of siblings is that the family circumstances for twins will typically be more similar than with siblings. More importantly, identical twins are genetically identical, whereas siblings, on average, share only half of their genetic endowments. Therefore, using data on twins, in particular genetically identical twins, may reduce the bias caused by heterogeneity within families.

We find large deleterious effects of conduct disorder on positive and negative human capital, even within pairs of identical twins. Conduct disorder increases the probability of grade retention and not completing high school. In addition, conduct disorder behaviours measured before the age of 18 have a strong effect on violent and criminal behaviour since the age of 18. For instance, within pairs of identical twins we find that conduct disorder reduces the probability of high school graduation with 4 to 13 percentage points and increases the probability of being arrested with 7 to 16 percentage points. Another important finding is that the effect of conduct disorder on human capital is more deleterious if these behaviours occur earlier in life. Various robustness checks, for instance sensitivity analyses that address the issues of measurement error and spill-over effects within pairs of twins, suggest that our main estimates may be lower bounds of the true effect of conduct disorder on human capital. We conclude that childhood mental health problems have high human and financial costs for families and society at large. Effective treatment early in life may yield high returns.

1.3 Software

Empirical results in this dissertation are produced using several software packages. Regression models (with and without ARMA errors) and the structural time series models (univariate and multivariate) in chapters 2 and 3 are estimated using the STAMP package, developed by Koopman et al. (2006). The intervention ARIMA models in chapter 2 are implemented using the STATA (2006) software. All microeconometric results in chapters 4, 5, and 6 are also produced using the Stata (2006) software. Graphics in this dissertation are created using the GiveWin interface of PcGive, developed by Hendry and Doornik (1996) and the Stata (2006) software. This dissertation is typeset in LaTeX using MikTeX and TeXnicCenter.

Chapter 2

Intervention Time Series Analysis of Crime Rates: The Impact of Sentence Reform in Virginia

This chapter is based on work done for my Tinbergen Institute M.Phil. thesis (2003), under the supervision of Prof. Siem Jan Koopman.

2.1 Introduction

The articles in the 2001 issue of the *Journal of Quantitative Criminology* (Volume 17, No. 4), with the notable contributions of Levitt (2001), Greenberg (2001), and Cantor and Land (2001), have prompted an interesting debate on methodologies for time series analysis of crime rates. In this chapter we would like to contribute to this discussion in the context of intervention analysis. Various intervention time series approaches have been used in the evaluation of programs and policies in a number of criminal justice settings (Loftin et al., 1983; McCleary and Hay, 1980; McDowall et al., 1980; McDowall et al., 1995). The standard approach to time series analysis in this framework aims at discriminating between the behaviour of the time series prior to the intervention and after the intervention (McCleary and Hay, 1980; Orwin, 1997). The typical research question is: "Given a known intervention, is there evidence that change in the series of the kind expected actually occurred, and, if so, what can be said of the nature and magnitude of the change?" (Box and Tiao, 1975). From the policy perspective it is important to assess whether a know intervention had the intended effect. For example, whether increased reliance on prisons, increased number of police, tougher gun control laws, and innovative criminal justice programs and policies reduced crime rates and deterred potential criminals from committing crimes.

A popular time series approach to intervention analysis is based on autoregressive integrated moving average (ARIMA) models as propagated by Box and Tiao (1975). Structural time series (STS) models provide another approach to the modelling of interventions. The STS approach is popularized by Harvey (1989) and has been applied in various policy and intervention analysis applications. For example, Harvey and Durbin

(1986) look at the effects on the number of car drivers killed in Britain after the seat belt legislation has been introduced in 1983. Harvey (1996) analyses the effects of the same British seat belt legislation using multivariate structural time series framework with control groups. Balkin and Keith Ord (2001) investigate the relationship between speed limit increases and traffic-related fatalities in the US. However, the STS approach has not been used extensively in crime analysis.[1]

In this chapter we adopt different time series methodologies to investigate the impact of parole[2] abolition and sentence reform in Virginia on reported crime rates. The Commonwealth of Virginia abolished parole and reformed sentencing for all felony[3] offences committed on or after January 1, 1995. This law was passed in a special legislative session in the autumn of 1994. Parole abolition was accompanied with substantially enhanced sentences for violent offenders.[4] To examine the impact of Virginia's abolition of parole on reported crime rates, we consider different empirical approaches to the intervention analysis. First we adopt regression methods to illustrate the problem of intervention analysis. Then, ARIMA models are applied as the standard approach to the intervention analysis of time series data, which are serially correlated, often nonstationary, and with strong seasonal and/or cyclical effects. Finally, the intervention analysis is carried out using the structural time series models, as an alternative to the ARIMA approach. Structural time series models are formulated in terms of components such as trends and seasonals, which have a direct interpretation and allow for time-varying parameters. The STS models are estimated both in the univariate and multivariate dimensions. We follow the approach of Harvey (1996), when multivariate STS models with control groups are introduced. Examined crime rate series include burglary, larceny, motor vehicle theft, robbery, aggravated assault, murder, and rape. Full details of the models and definitions of the crime categories are given in the chapter.

The focus of this chapter is on the impact of the new legislation on reported crime rates. Policy changes that increase the expected punishment per crime can lead to both greater deterrence and greater incapacitation (Kessler and Levitt, 1999). By looking at changes in crime immediately after the introduction of a sentence reform in Virginia, we hope to isolate a pure deterrent effect of the new legislation that is not contaminated by the effect of incapacitation. Hence, to the extent that severity of punishment serves as a deterrent to committing crimes in the short run, we would expect the reported crimes to drop especially for the violent offences. Given a change in the Virginia's legislation, we want to test if there was a decline in the analysed crime rates and what was the magnitude of this decline.

[1] To our knowledge, the structural time series methodology applied to crime data is carried out by Harvey and Fernandez (1989) and Atkinson et al. (1997), who look at the number of outliers and breaks in the monthly number of purse (handbag) snatches in Hyde Park in Chicago. Koopman et al. (2008) model recidivism behaviour of juveniles from a Dutch Judicial Juvenile Institution, using a non-Gaussian STS model. Finally, Vujić and Koopman (2008) (chapter 3 of this dissertation) investigate the cyclical behaviour of annual recorded burglary data in the UK.

[2] Parole is the releasing of a prisoner either temporarily or before his/her period in prison is finished, with the agreement that he/she behaves well; "good-time" credit; "earned sentence" credit.

[3] Felony is (an example of) a serious crime which can be punished by one or more years in prison.

[4] Virginia Criminal Sentencing Commission (1995), Annual Report, Virginia Criminal Sentencing Commission, Richmond, VA.

Our sample includes the 1990 to 1999 period when considerable social and economic changes occurred in the United States. There were declines in crime trends throughout the US during this decade. Furthermore, the second half of the 1990s was an economically prosperous period in the US. For example, unemployment rates declined sharply through most of this period. It was also a period in which a number of innovative criminal justice programs and policies were enacted both at the State level and at the local communities level. Favourable changes in patterns of drug use and access to guns were put in place. These factors could serve as alternative explanations for the decline in crime throughout the US in general, and Virginia in particular. Disentangling the impact of parole abolition on crime rates in Virginia from these other factors poses a considerable methodological challenge.

The findings of this chapter are that crimes in the eligible categories responded immediately to the passage of the new legislation. We see not only an immediate fall in crime with the adoption of the new legislation, but also a continued decline in crime rates over several years after the passage of the new legislation. We could say that the deterrence (immediate) effect of the new legislation was about -0.1% and over the longer run as incapacitation and further deterrence kicked in, the effect of the new legislation was around -7%. When it comes to the effect of the new legislation on the particular crime offences, we find that the new legislation significantly affected burglary (-4 to -5) and murder series (-0.10). When it comes to rape, we find a negative significant result (-0.24) only in one of our bivariate model specifications, which is not really superior to the univariate estimation result (-0.15), found not to be significantly different from zero. We do not find that the new legislation significantly affected the aggravated assault which is, by definition, a violent offence.

The remaining part of the chapter is organised as follows. In Section 2.2 we discuss in more details the criminal justice situation in Virginia and its recent changes and developments in the parole and sentence systems. Different time series methodologies for intervention analysis are considered in Section 2.3. Additionally, this section gives plan and details of empirical intervention study, in particular using descriptive and regression approaches. Empirical results of our investigation of the effect of parole abolition and sentence reform on the crime rates in Virginia, using ARIMA and structural time series methods, are presented in Sections 2.4, 2.5, and 2.6. Section 2.7 offers discussion of the results, comparing different methodological approaches to intervention analysis. This section also concludes and raises questions that can be further analysed.

2.2 Changes in criminal justice system of Virginia

The abolition of parole in Virginia was proposed during the 1993 George Allen's campaign for a Governor. A key element of the campaign was to reduce the disparity between the sentence imposed in court and the actual time-served. This meant to eliminate or reduce "good-time" credit and abolish parole. As a Governor, Allen established the Commission on Parole Abolition and Sentencing Reform. This Commission formed by crime victims, law enforcement professionals, judges, prosecutors, business and civil letters, and other state and local officials recommended a "plan to abolish parole, establish truth-in-sentencing, incarcerate violent and repeat offenders significantly longer, institute more productive and economical methods to punish non-violent criminals, and expand prison capacity."[5]

In September 1994, a special session of the Virginia General Assembly was held to take up the recommendations of the Governor's Commission. After days of deliberation and compromise, parole was abolished for offenders convicted of a felony committed on or after January 1, 1995. This initiative abolished parole, established a guidelines-based truth-in-sentencing system, and increased sentence length for violent offenders. We shall briefly summarize the changes in the sentencing reform introduced in Virginia after January 1, 1995:

i. **Parole abolition and "truth-in-sentencing".** Under the previous system, as specified in Code of Virginia §§53.1-202.2 and §§53.1-202.3, a combination of earned good-time credits and parole could allow a convicted offender to serve a fraction of the imposed sentence. Under the new system, offenders of crimes committed on or after January 1, 1995 can earn a maximum 15% reduction in their time served, and must serve at least 85% of their imposed sentence. It was also envisioned that an offender would have to earn the reduction by participating and cooperating in programs in which they were assigned. Inmates who were sentenced according to the new parole system would serve closer to 90% of their sentences. This collection of different but related publice policies aimed to abolish parole, so that convicts serve the period that they have been sentenced to, was also known as "truth-in-sentencing" (TIS). For example, in 1994, a federal TIS law was passed. To qualify for TIS federal funding, offenders must serve 85% of their sentence for qualifying crimes.

ii. **Time served for first-time violent offenders was increased by 100% to 125%.** In addition to a significant reduction in "good-time" credit, the enabling legislation established criteria for developing discretionary felony guidelines. The newly created Sentencing Commission was required by §17.237 of the Code of Virginia (re-codified to §17.1-805) to adopt sentencing guideline midpoints based on actual time-served "for similarly situated offenders, in terms of their conviction offence and prior criminal history, increased by 13.4 percent" (to account for earned-sentence credits) "and second, by eliminating the upper and lower quartiles".[6] To address offenders who committed violent offences, the sentence for first and second degree murder, forcible rape, forcible sodomy, object sexual penetration and aggravated sexual battery were increased by 125%. For voluntary manslaughter, robbery, aggravated malicious wounding, malicious wounding and any burglary of a dwelling or any burglary with a firearm the midpoint was increased by 100% over historical time-served. These increases were established for offenders with no previous violent offence conviction. Hence, the time served for first-time violent offenders was increased by 100% to 125%.

iii. **Time served for repeat violent offenders was increased by 300% to 500%.**

[5] Governor's Commission on Parole Abolition and Sentencing Reform, Final Report, August 1994.
[6] Code of Virginia.

> Offenders with prior convictions for violent felonies were also recommended for additional enhancements over historical time-served. The General Assembly established by statute the offences that would constitute violent offences. As expected murder, rape, robbery and felony assault charges were included. In addition some offences not usually thought of as violent (e.g., burglary, obscenity and possession of weapons offences) would increase the sentencing guidelines recommendation.[7] An offender with a prior conviction for a violent offence with a statutory maximum of forty years or more would have a guidelines recommendation increased by 300% to 500% over historical time-served. A prior violent felony with a statutory maximum of less than forty years increased the guidelines recommendation by 100% to 300%.[8]

The net result of the implementation of the legislation was a substantial increase in the sentences for the violent offences (especially rape and murder) and also for offenders with a violent past. Table 2.1 (adapted from the Virginia Criminal Sentencing Commission annual report of 1995) compares the median time-served (in years) for prisoners released in 1993 (in a system with parole) with a median expected time-served for two groups of offenders sentenced in 2001 a system without parole. Three groups of offenders sentenced in 2001 are described in Table 2.1: (a) group of offenders who did not have any prior offences; (b) group that had prior offences with a statutory maximum less than 40 years (roughly corresponding to non-violent prior offence); (c) group of offenders that had prior offences with a statutory maximum greater then 40 years (roughly corresponding to a prior record with violent offences).

The new sentence reform incorporates guidelines with significant increases in recommended prison sentences for all violent offenders. As can be seen from Table 2.1, increases in time-served were especially high after the implementation of the legislation for robbery, aggravated assault, murder, and rape.

Another example of the time served under the new truth-in-sentencing system and old parole system is presented in Figure 2.1.[11] In Figure 2.1, prison time served under the

[7] See Code of Virginia §17.1-805(C) or Sentencing Guidelines Manual 7th Edition Table A.

[8] Virginia Criminal Sentencing Commission, 1995 Annual Report, page 10-11 Note: First and second degree murder, rape, forcible sodomy, object sexual penetration and aggravated sexual battery midpoints were increased by 125% for no prior violent felony, 300% for prior violent felony with statutory maximum less than forty years and 500% with prior violent felony with a maximum of 40 years of more. A first degree murder conviction in combination with a prior violent felony with a maximum of forty years or more is recommend for a life sentence. Voluntary manslaughter, robbery, aggravated malicious wounding, malicious wounding, burglary of a dwelling, burglary with a weapon midpoints were increased by 100% for no prior violent offence, 300% for prior violent felony with less than forty year maximum, and 500% with prior violent felony with forty or more maximum. Sale etc. of a Schedule I/II drug midpoints were increased only for offenders with a prior violent felony; a 200% increase for prior felony with maximum of less than 40 years and 400% increase for violent felonies with maximum of 40 years or more. Midpoints were increased for all other felonies by 100% for prior felony with maximum of less than forty and 300% for prior felony with a maximum of forty years or more.

[9] FY93 Used because parole was an issue in the 1994 campaign and parole grant rates began to change prior to the abolition of parole.

[10] Virginia Criminal Sentencing Commission Annual Report 1995 p7 for FY93 Actual Time Served and Annual Report 2001, pp. 66-71. Burglary, Motor Vehicle Theft and all combined data is from unpublished data maintained by the Sentencing Commission.

[11] Adapted from "Crime in the Commonwealth, 1988-1998", web address:

Table 2.1: Comparison of median time-served (in years) in 1993 (system with parole) and anticipated median time-served for offenders sentenced in 2001 (system without parole)

Offence	Released FY93[9] Median time	Sentenced FY01 Median expected time			
		No prior	Category II Prior < 40	Category I Prior ≥ 40	All combined
Burglary	2.2	1.8	3.6	5.4	2.7
Larceny	1.3	1.1	1.8	2.3	1.4
Motor vehicle theft	1.3	1.3	1.8	2.7	1.4
Robbery	**4.4**	6.4	11	16.2	**7.3**
Aggravated assault	**2.8**	3.7	6.2	7.3	**4.1**
Murder (2nd degree)	**5.7**	13.6	22.7	20.0	**16.3**
Rape (forcible)	**4.4**	9.0	13.5	34.3	**12.6**

Source: The Virginia Criminal Sentencing Commission annual report, 1995

parole system is compared to time served under truth-in-sentencing for offenders convicted of first degree murder, forcible rape, and robbery with firearm. Parole system time served is based on time served by inmates released from prison from 1988 to 1992. Truth-in-sentencing time served is estimated based on sentence length for cases sentenced in 1998. All sentence lengths shown are median values.

The upper part of Figure 2.1 shows that under the previous parole system, offenders convicted of first degree murder with no prior violent record served 12.4 years in prison, whereas under the truth-in-sentencing system, offenders convicted of this offence would now serve more than 37 years in prison. Offenders with a Category II record who were serving about 14 years in prison will now serve 51 years. Offenders with more serious Category I records who were serving about 15 years will now serve more than 95 years in prison. Analogous interpretations are for rape and murder convictions. Offenders convicted of forcible rape and armed robbery will receive much longer sentences as a results of the new reform. Similar increases have occurred in time served for offenders convicted of other violent offences, as well as for other property and drug offenders with violent prior records.

During the 1990s, Virginia law makers have enacted various laws to respond to the rising crime in Virginia. Table 2.2 summarizes some of the main initiatives passed during this period. In our opinion, only Virginia firearms transaction programme in 1989, one handgun per 30-day purchase limit in 1993, and parole abolition and truth-in-sentencing in 1994 could have served as a deterrent to potential criminals.

To the extent that severity of punishment serves as a deterrent to committing crimes, we would expect that the Virginia's anti-crime initiatives reduce the reported crime rates, especially for violent offences. Figures 2.2 and 2.3 in Section 2.3 show a sharp decline in the analysed crime series during the 1990s, indicating that these anti-crime initiatives

`http://www.dcjs.state.va.us`

Table 2.2: Major Criminal Justice Initiatives in Virginia, 1988-1998

Year	Initiative	Description
1989	**Virginia firearms transaction programme**	A criminal history records check to be conducted on persons purchasing a firearm from a licensed dealer.
1990	DNA analysis & data bank	Persons convicted of a felony must provide a blood sample to produce a DNA profile for storage in the DNA data bank.
1993	**One handgun per 30-day purchase limit**	Limits to one the number of handguns that may be purchased in any 30-day period.
	Serious or habitual offender comprehensive action programme (SHCOAP)	City and county governments can establish multi-agency SHOACPs to share information about serious juvenile offenders.
	Juvenile criminal history records	Police records should maintain fingerprints & case disposition information for juveniles age 13 and older charged with a felony.
1994	Sex offender registry	Police should maintain a registry of persons convicted for sex offences against minors.
	Parole abolition and truth-in-sentencing	Policy intervention which effect we want to empirically test.
	Community-based correction system for state-responsible offenders	Community-based alternative sanctions for state-responsible-offenders.
	Community correction act for local responsible offenders	Community-based corrections programmes as sentencing alternatives.
	Pre-trial services act	Localities can operate pre-trial services programmes to assist judicial officers in bail-related duties.
1995	Virginia juvenile community crime control act	Community-based system of progressively intensive sanctions and services corresponding to the severity of offence, treatment needs and crime trends in the localities.
	Crime victim and witness rights act	Provides crime victims and witnesses with certain legal rights.
1996	Juvenile justice reform	In juvenile proceedings, the welfare of the child and family, community safety, and victims' rights are of paramount concern.

Source: Crime in the Commonwealth, 1988-1998

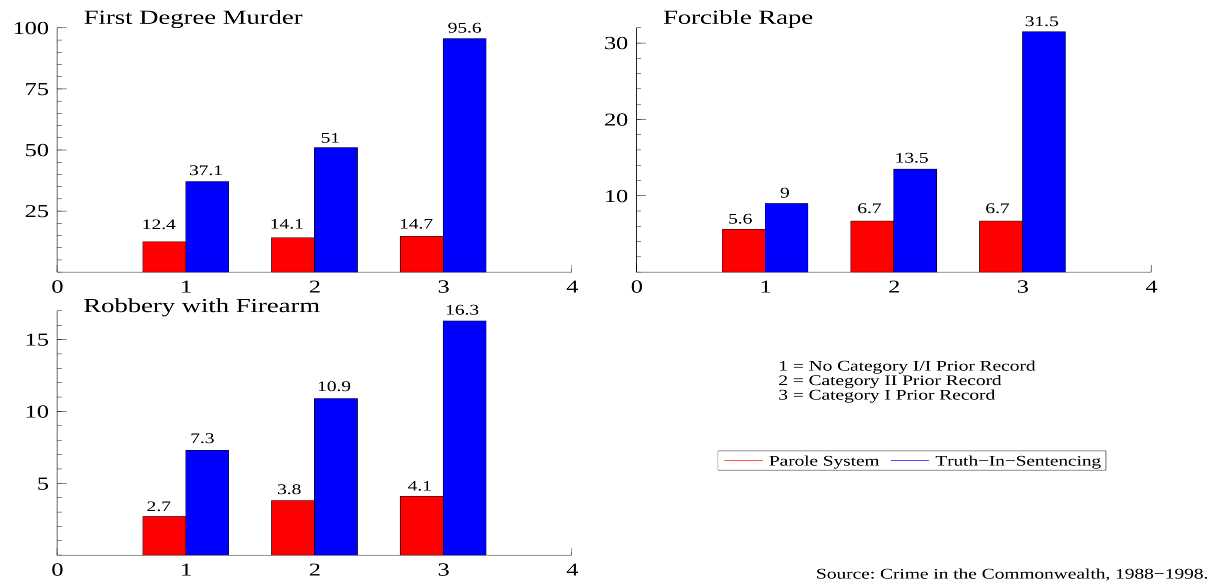

Figure 2.1: Time served in prison under parole system and truth-in-sentencing system

might have had the intended effect. It should be noted that sharp declines in violent crime rates have also occurred nationwide and severity of punishment is only one explanation for a drop in crime. A number of alternative explanations can also be used to explain a drop in crime. A recent book, "The Crime Drop in America" (Blumstein and Wallman, 2000) compiles a variety of explanations for the reductions in crime in the US in the 1990s. For example, alternative explanations for drops in crime include: changes in drug use patterns (Johnson et al., 2000), policing and community policing (Eck and Maguire, 2000), growth in prison expansion (Spelman, 2000), reductions in use of handguns (Blumstein and Wallman, 2000), expanding economy (Grogger, 2000), and changing demographics (Fox, 2000). According to Levitt and Dubner (2005), the most important crime-drop explanations are increased reliance on prisons, increasing the number of police officers per capita, the crash if the crack market, and the legalisation of abortion.

Despite the complexities inherent in understanding the factors associated with the declining crime rates, Virginia's experience with abolition of parole and sentence reform remains of interest for a few reasons. A number of States have abolished parole for specific felony offences, while Virginia abolished parole for all felony offences. Parole abolition was further accompanied by large-scale changes in the sentencing system. Further, the timing of this law occurred when the downward trends in crime had already begun both nationwide and in Virginia. It is therefore interesting to empirically investigate whether parole abolition and sentence reform in Virginia led to steeper declines in crime as compared to expected patterns based on historical data.

2.3 Plan and details of empirical intervention study

2.3.1 Data description

The data analysed in this chapter come from the Uniform Crime Reports (UCR) collected by the Virginia State Police, in the period 1984 to 2006. The pre-intervention period corresponds to the period between January 1984 to December 1994. The post-intervention period corresponds to the period between January 1995 to December 2006. Analysed crime series are burglary, larceny, motor vehicle theft, robbery, aggravated assault, murder, and rape. In order to obtain an insight in our crime data set, we shall give a brief overview of the Virginia crime reporting practices.

"UCR is a city, county, state, and Federal law enforcement programme that provides a standardized, nationwide view of crime based on data submitted by law enforcement agencies throughout the country." Hence, we have constructed the data recorded according to the UCR approach, because it enables comparisons of crime statistics across all the US. Uniform Crime Reports summary data consist of aggregate numbers of offences known to police and arrests. Offences are "the number and kind of criminal acts that are reported to the police." The offence rate is, therefore, an indicator of criminal victimization. Arrests are "a measure of police activity as it relates to crime," Roberts (2005).

UCR summary offences are classified into two groups: Category I and Category II. Category I offences are any prior conviction or juvenile adjudication for a violent crime with a statutory maximum penalty of 40 years or more. Category I offences include: criminal homicide, forcible rape, robbery, aggravated assault, burglary, larceny motor vehicle theft, arson. Category II offences are any prior conviction or juvenile adjudication for a violent crime with a statutory maximum penalty less than 40 years. Category II offences include: other assaults, forgery and counterfeiting, false pretences/swindle/confidence games, embezzlement, stolen property offences, destruction/damage/vandalism of property, weapon law violations, prostitution and commercialized vice, sex offences (except rape and prostitution), narcotic drug laws, gambling, offences against the family, driving under the influence, liquor law violations, public drunkenness, disorderly conduct, all other offences (except traffic), curfew/loitering, runaway, juvenile. Category II offences are recorded in the UCR system only if an arrest occurred. In this chapter we shall analyse Category I offences (except arson), because they are most serious and/or most frequently reported offences and are "the best indicator of crime," Roberts (2005). Definitions of these crime categories are given in the Appendix 2.A.2.

The UCR approach is called a "summary" approach because it reports only the most serious offence in a criminal incident, following the Hierarchy Rule. According to the Hierarchy Rule, the most-to-least serious offences are: criminal homicide, forcible rape, robbery, aggravated assault, burglary, larceny, and motor vehicle theft (arson is not subject to this rule). Starting from 2000, Virginia also reported crime data in incident-based format (IBR), where all offences associated with a criminal incident are reported. For example, a homicide that occurs during a robbery with a firearm would be counted as a homicide under the UCR system (one criminal offence), while under the IBR system all three offences would be captured (homicide, robbery, and weapon law violation). Although the IBR approach to recording crime gives a much more detailed picture of crime than the summary system, there are only six States that report crime in the incident-based format. Further, the total number of IBR offences is no more than 3.61% higher than the total number for the same offences counted in the UCR system. For the purposes of this chapter, we shall use the data in the UCR format in order to ensure consistency of the data in the pre- and post-intervention periods. For more details on UCR versus IBR crime reporting in Virginia, see Roberts (2005).

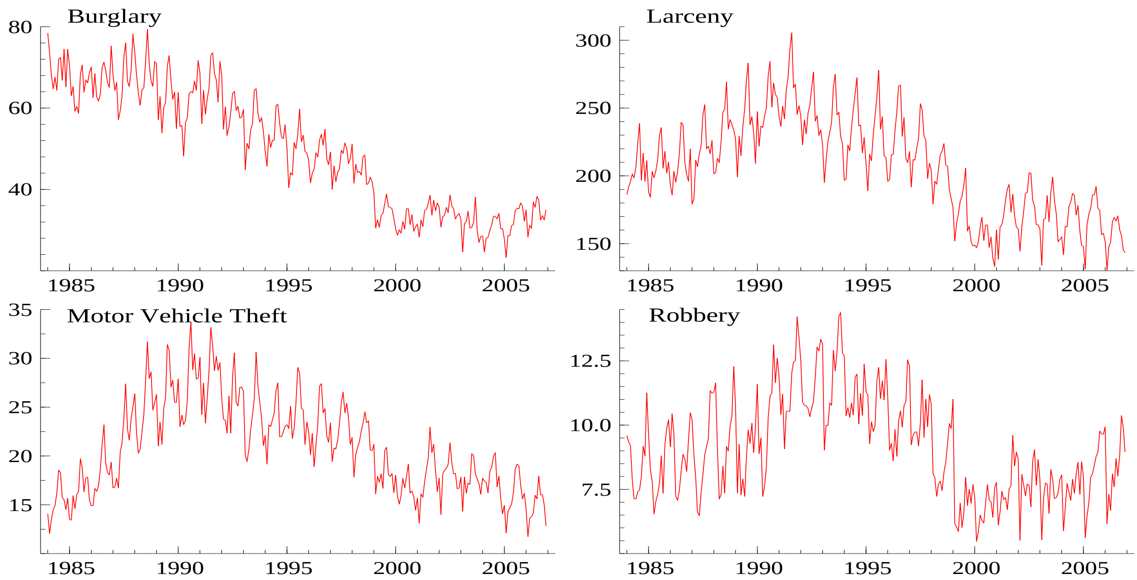

Figure 2.2: Property crime rates: burglary, larceny, motor vehicle theft, and robbery.

Figures 2.2 and 2.3 present the reported crimes rates per 100,000 population, for property and violent crimes respectively.[12] It can be observed from Figures 2.2 and 2.3 that most of the crime series were declining at about the same time that Virginia enacted major legislative initiatives to reduce violent crime. "Research to date is unable to determine if these reductions in crime rates are due to specific anti-crime initiatives."[13] Reductions have occurred in the types of crimes that were targeted by these initiatives, indicating that they may be having their intended effect. Declines in violent crime rates in Virginia coincided with declines in violent crime rates nationwide. There is still a large debate among criminologists and policy makers as to which factors contributed to the crime-drop in the US (e.g., legalized abortion 20 years earlier, expanding economy, community policing, changes in crack and other drug markets, and/or higher arrest and incarceration rates). This chapter tries to contribute to better understanding of the statistical relationship between anti-crime efforts and crime reductions over time.

Policy changes that increase the expected punishment per crime can lead to both greater deterrence and greater incapacitation. Empirical evidence which links increased punishment with lower crime rates is consistent.[14] According to Levitt and Dubner (2005),

---

[12] Following the UCR categorization scheme, robberies were included together with the property crimes.

[13] Crime in the Commonwealth, 1988-1998.

[14] See for example Ehrlich (1973), Grogger (1991), Kessler and Levitt (1999), Levitt (1997), Marvell and Moody (1994) or Marvell and Moody (1996).

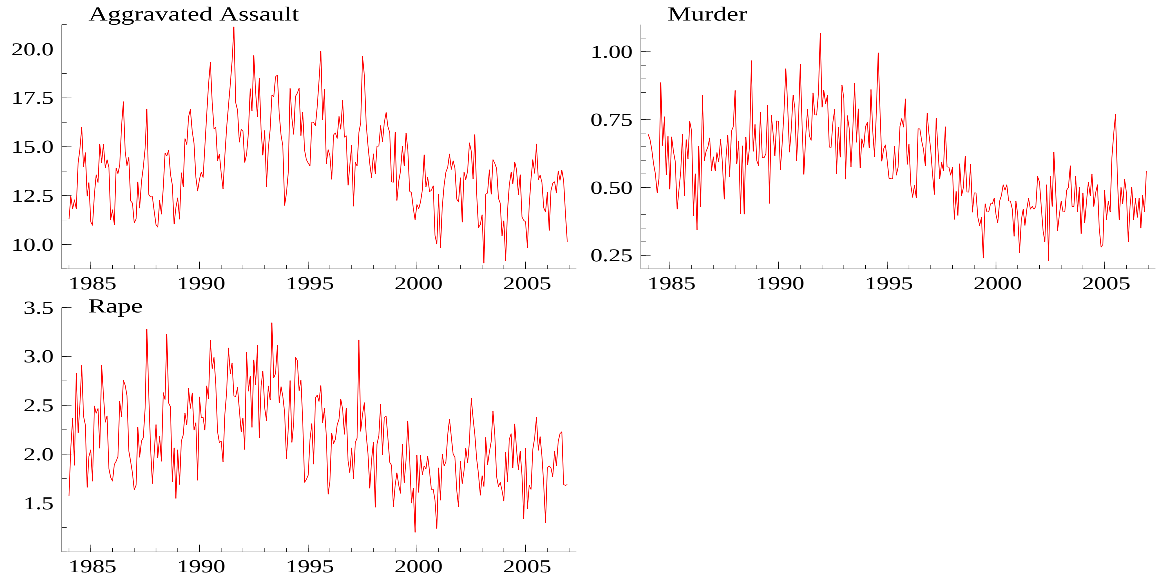

Figure 2.3: Violent crimes rates: aggravated assault, murder, and rape.

increases in prison population account for roughly one-third of the drop in crime in the US. However, most empirical tests on deterrence do not separate the effect of deterrence from the effect of incapacitation. Short-run declines in crime are likely to be attributable to deterrence, whereas the incapacitation effect of sentence enhancements will occur only in the long-run (Kessler and Levitt, 1999). In case of Virginia, the 1994 legislation abolishing parole and establishing a truth-in-sentencing system was a single, most significant factor affecting the size of prison population. Although it took time for the longer prison sentences imposed under the 1994 sentencing reform to have a significant growth effect on Virginia's prison population, decrease in the parole grant rate had an almost immediate effect on the size of the prison population.[15] By looking at changes in crime immediately after the introduction of a sentence reform in Virginia, we hope to isolate a pure deterrent effect of the new legislation that is not contaminated by the effect of incapacitation. Hence, to the extent that severity of punishment serves as a deterrent to committing crimes in the short run, we would expect the reported crimes to drop especially for the violent offences. Given a change in the Virginia's legislation, we test if there was a decline in the analysed crime rates immediately after the introduction of a new law and what was the magnitude of this decline.

[15]For example, between 1990 and 1993, Virginia's annual parole grant rate averaged about 41% (i.e., about four out of ten prisoners elgible for parole were granted parole). The parole grant rate begain to decline in 1993, and by the end of 1994 it dropped to about 14%. After the parole system was abolished in 1994, the grant rate remained below 20% (Crime in the Commonwealth, 1988-1998). Further, sentencing reform applied to virtually all felony convictions, while repeated violent offenders had to spend from two to more than five times longer in prison than under the parole system.

2.3.2 Intervention analysis

Examples of intervention effects in time series are given in a graphical form in Figure 2.4. The first graph is a so-called *pulse intervention* and is used to capture a single special event in a month such as a special holiday or a strike. Such events may cause outlying observations within the time series and the pulse regression variable can take such observations outside the general model. The pulse intervention variable at time τ is defined by

$$\begin{aligned} I_t &= 0, \qquad t < \tau, \qquad t > \tau, \\ &= 1, \qquad t = \tau. \end{aligned} \tag{2.1}$$

The second graph shows a so-called *step intervention* that enables breaking the single time series into two distinct segments with two different overall means, one consisting of all pre-intervention observations and one consisting of all post-intervention observations. The step intervention is introduced in the model to capture events such as the introduction of new policy measures or changes in regulations. The step intervention variable at time τ is defined by

$$\begin{aligned} B_t &= 0, \qquad t < \tau, \\ &= 1, \qquad t \geq \tau. \end{aligned} \tag{2.2}$$

Abrupt break interventions such as B_t may not occur regularly in practice. The effect of a policy change is usually not felt instantaneously. The change to a new mean level may come into effect gradually during a particular period. We do not want to rule out such interventions and therefore we also consider the *smooth break intervention* as given by

$$\begin{aligned} S_t &= 0, && t \leq \tau_1, \\ &= (t - \tau_1) \;/\; (\tau_2 - \tau_1), && \tau_1 < t \leq \tau_2, \\ &= 1, && t > \tau_2. \end{aligned} \tag{2.3}$$

It is apparent from the third graph of Figure 2.4 that the intervention starts to take effect from time period τ_1 while it takes full effect after a period of length $\tau_2 - \tau_1$.

The analysis of interventions in a time series focuses on a test of the null hypothesis, that is, did the intervention have an impact on the time series? If so, what can be said of the nature and magnitude of the change? In case of the step interventions B_t and S_t the null hypothesis can be tested by comparing means of the pre- and post-intervention parts of the time series.

We first model the intervention as the level shift or step intervention B_t, where the value of the level of the time series instantly changes at the time point when the intervention takes place, and where the level change is permanent after the intervention. The effect of this immediate change is modelled as a level break, so that a step change in the input produces a "ramp" (gradual) response in the output (Box and Tiao, 1975). The impact of the intervention is assumed to start instantaneously from January 1995, since Governor Allen was elected in November 1993 on a platform whose primary focus was parole abolition and sentence reform. Even before parole was formally abolished, there was a considerable focus in the media on the "severity" of the new parole abolition system

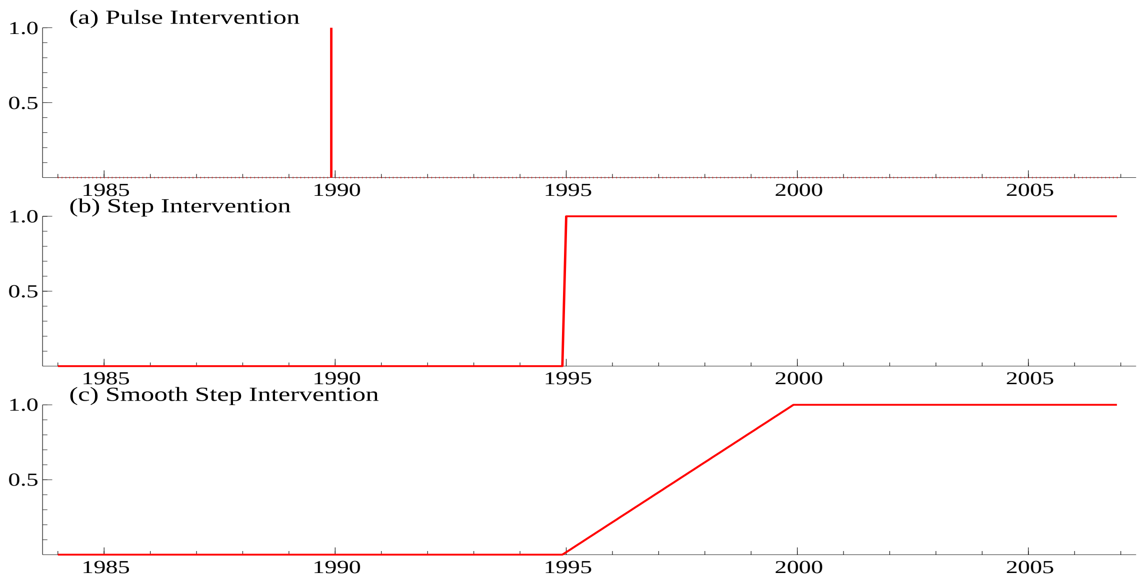

Figure 2.4: Intervention effects: (a) pulse intervention, modelled as an additive outlier; (b) step intervention, modelled as a level break; (c) smooth step intervention, modelled as a gradual level break.

and the enhancements in time-served for violent offenders. Although it took time for the longer prison sentences imposed under the 1994 sentencing reform to have a significant growth effect on Virginia's prison population[16], decrease in the parole grant rate had an almost immediate effect on the size of the prison population.[17] Hence we first assume that the deterrent impact of the new legislation have occured very soon after January 1st, 1995. Later, as part of a robustness analysis, we will investigate whether the impact had a more gradual effect on the recorded crimes.

In order to empirically test the hypothesis, the intervention analysis of the impact of parole abolition and sentence reform on recorded crime rates is assessed through various time series approaches. The effects of the intervention are estimated using regression, ARIMA, and structural time series models. In all three different models, the intervention is introduced as a step function. The results do not lead to conclusive evidence and therefore more elaborate analyses are considered in Sections 2.5 and 2.6. However, by taking the various modelling approaches of this section and the Sections 2.4 and 2.5 together, we conclude that some deterrent impact of the new legislation is noticeable for crime series with a violent nature.

[16]The longer prison sentences imposed under the 1994 sentencing reform could have had a significant growth effect on Virginia's prison population only from about year 2000 and after.

[17]At the beginning of 1995, the inmates confined for offences committed before January 1, 1995 were still admitted under the old parole system. However, in early 1996, only about 25% of its new inmates admitted to the prison came in under the old parole system (by the end of 2000, this number was about 1%).

2.3.3 Preliminary data analysis

Before we investigate the time series and the intervention effects on the basis of statistical models, we first compare the changes in the means of the entire pre-intervention period (1984-1994) with the post-intervention period (1995-2006). In Figures 2.5 and 2.6, the two different means are presented for the property crime series and the violent crime series before and after the intervention. It appears that both property crime rates and violent crime rates are affected by the new legislation; decreases in all recorded crimes are apparent after January 1995. Since this analysis considers a fairly long pre- and post-intervention period that potentially corresponds to multiple temporal regimes, a better understanding of the change in crime rates may be obtained by restricting the sample to four years before the introduction of the law and four years after the introduction of the law. Performing the same analysis for the sub-sample from January 1991 to December 1998 and comparing the means of the analysed crime series before and after the intervention, still indicates a drop in all recorded crimes after January 1995 (the graphs not shown here).

These results may provide a misleading picture of the change, because no information on trends are incorporated in the calculations. When fixed trends are considered, the differences between the means of the pre- and post-intervention remain very similar to the previous analysis without a trend, see Figures 2.7 and 2.8. All series show a drop after January 1995. However, for the sub-sample from January 1991 to December 1998, we draw a different conclusion (Figures 2.9 and 2.10). It seems that new legislation had as a result a drop in burglary, murder, and rape crimes, an increase in larceny and motor vehicle theft crime (substitution effect), and had almost no effect on aggravated assault. This preliminary analysis shows that measurements of the intervention effect can be rather different and the need for an elaborate analysis based on time series models becomes imminent. In the next section, we shall perform a Student's t-test for testing for a change in mean of the analysed series and discuss the appropriateness of the fitted models.

2.4 Preliminary approaches to intervention analysis

2.4.1 Regression analysis

In a classical regression analysis the impact of an intervention is modelled by

$$y_t = x_t'\beta + \delta B_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \tag{2.4}$$

for $t = 1, \ldots, n$, where y_t is a time series of crime rates, x_t is a $k \times 1$ vector of explanatory variables (or covariates) and β is a $k \times 1$ vector of regression coefficients. Constant, trend and seasonal dummies can be included in the vector of covariates x_t, together with other explanatory variables that may have an influence on crime rates. The variable that measures the intervention effect is defined as a dummy variable B_t which equals zero before a fixed time point and equals to one on and after this time point, see the definition in (2.3). In our empirical study, the intervention variable B_t is zero for the period before January 1995 and is coded to one for the period on and after January 1995.

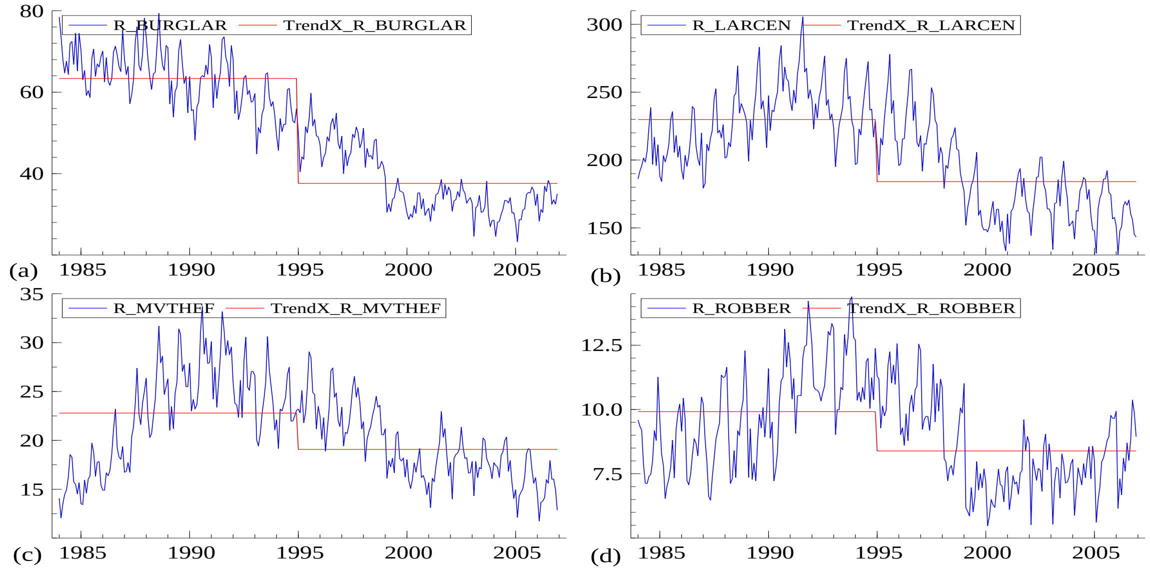

Figure 2.5: Mean change in property crime rates: (a) burglary; (b) larceny; (c) motor vehicle theft; (d) robbery.

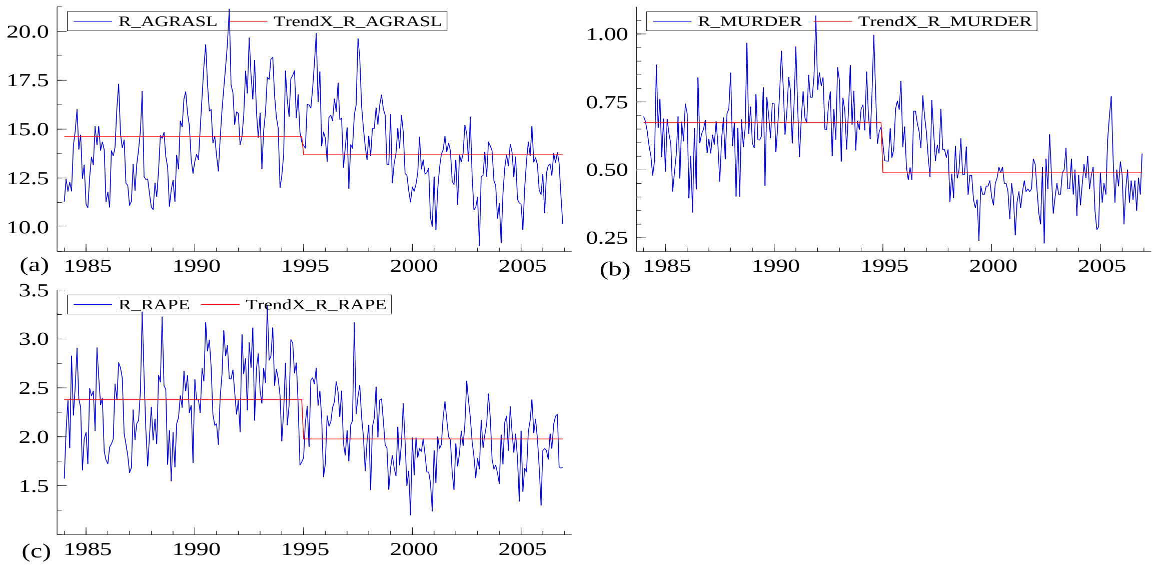

Figure 2.6: Mean change in violent crime rates: (a) aggravated assault; (b) murder; (c) rape.

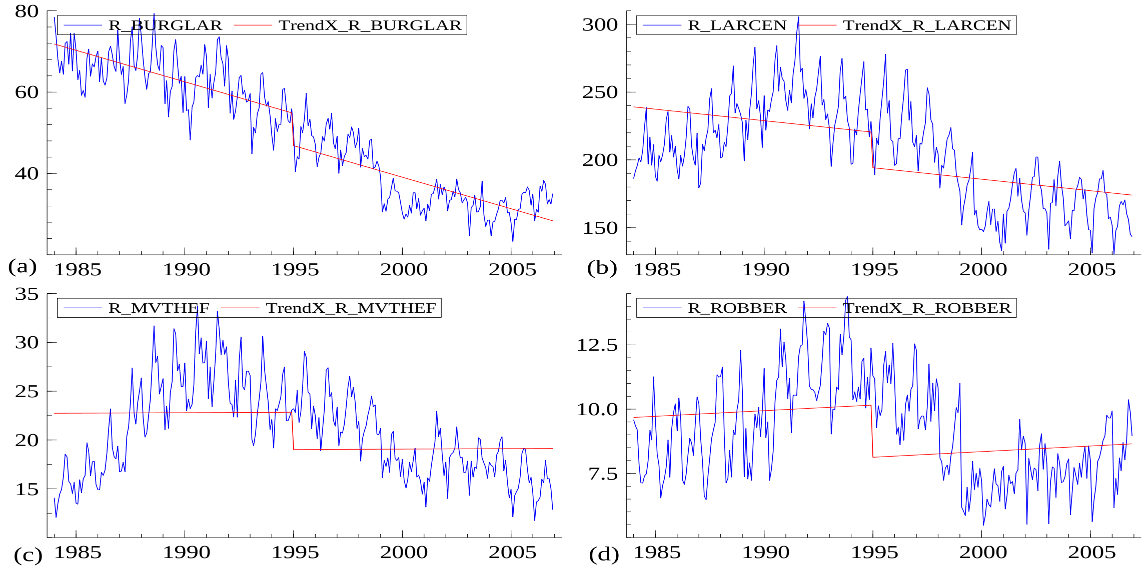

Figure 2.7: Trend change in property crime rates: (a) burglary; (b) larceny; (c) motor vehicle theft; (d) robbery.

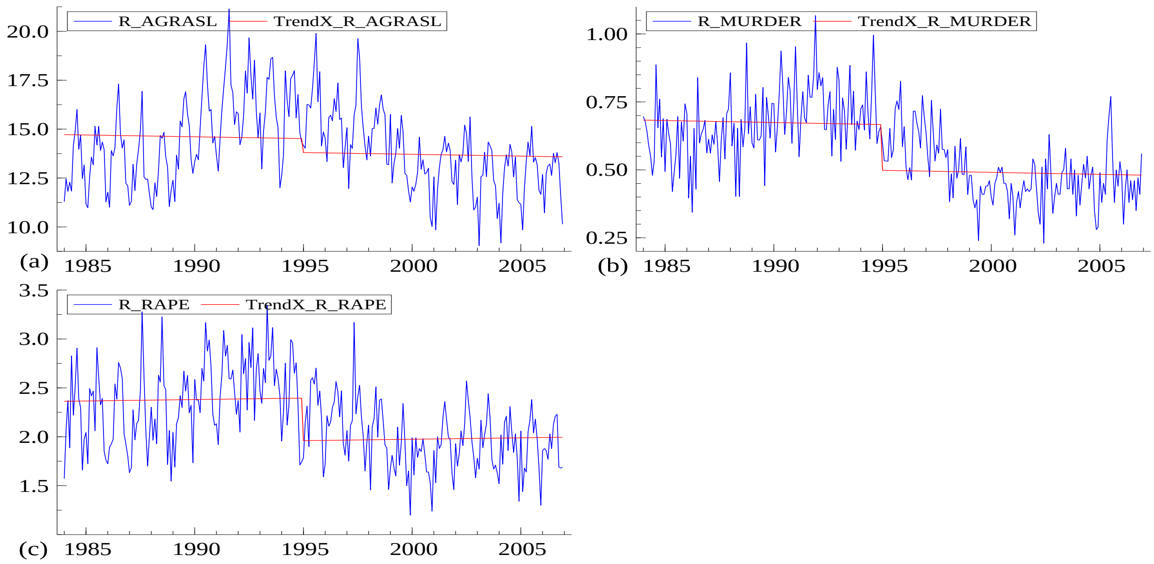

Figure 2.8: Trend change in violent crime rates: (a) aggravated assault; (b) murder; (c) rape.

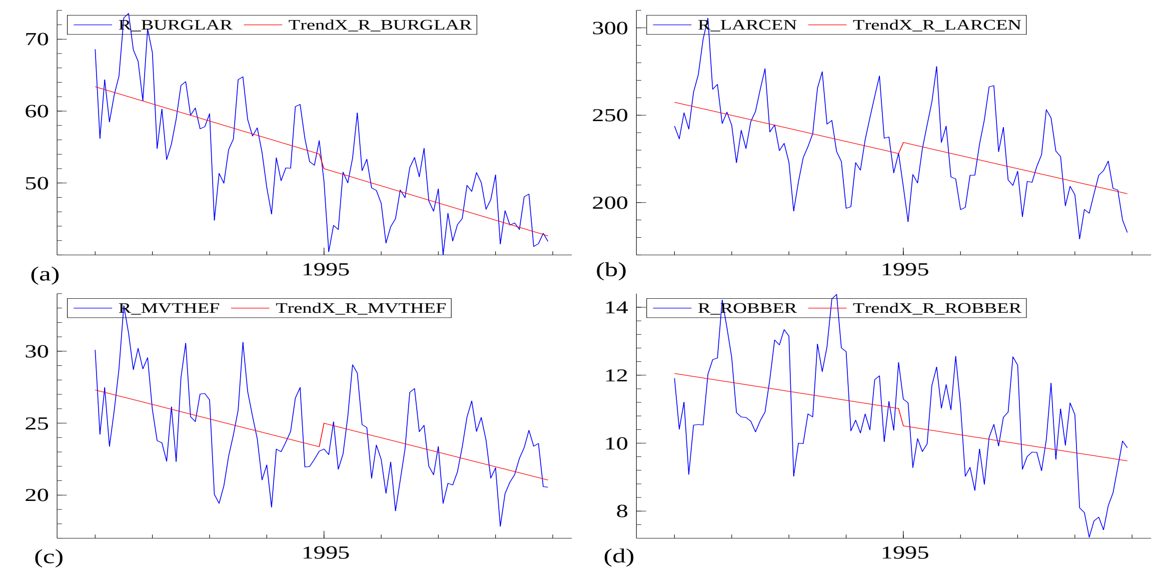

Figure 2.9: Trend change in property crime rates. Subsample 1991:01-1998:12: (a) burglary; (b) larceny; (c) motor vehicle theft; (d) robbery.

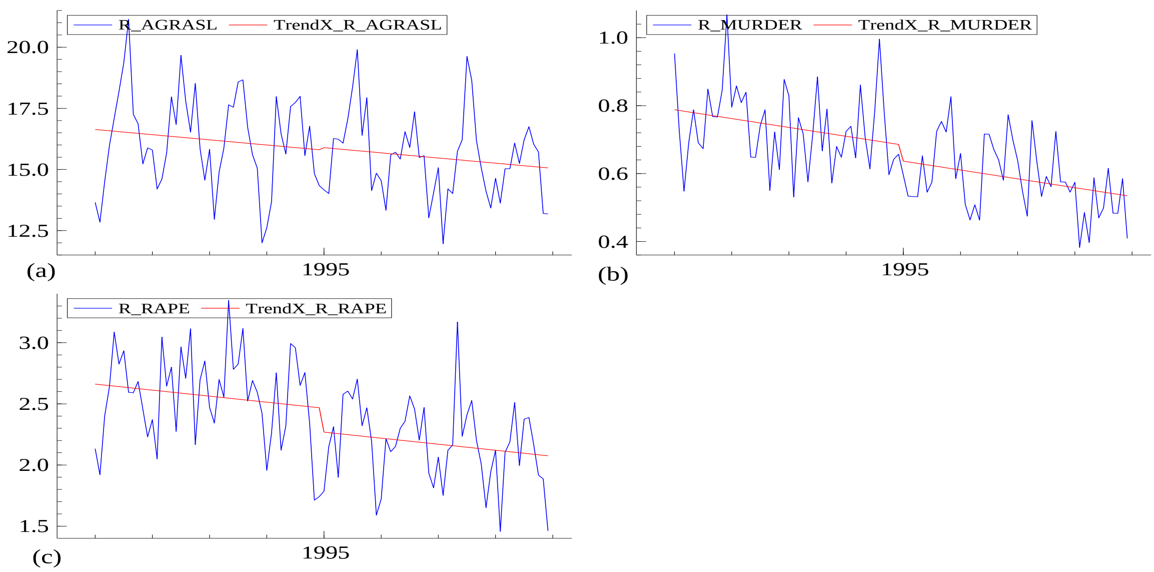

Figure 2.10: Trend change in violent crime rates. Subsample 1991:01-1998:12: (a) aggravated assault; (b) murder; (c) rape.

The intervention coefficient δ measures the change in the mean of crime rates time series before and after the intervention period. For the regression model (2.4), the method of ordinary least squares (OLS) can be used to estimate β and δ simultaneously.

The estimator δ depends on the model assumed for the disturbance. If there are a large number of observations before and after the intervention, then δ is estimated as the difference in the two means, though its variance depends on the properties of the disturbances ε_t. However, Student's t-test for estimating and testing for a change in mean is valid "only if the observations before and after the event of interest varied about means μ_1 and μ_2, not only normally and with constant variance but independently," see Box and Tiao (1975). Because our data are in the form of time series, serially dependent, nonstationary and with strong seasonal effects, we cannot assume that the disturbances ε_t are normally and independently distributed with mean zero and variance σ_ε^2 for all time points $t = 1, \ldots, n$. Therefore, available Student's t-test would not be a valid approach for testing for a change in mean before and after the intervention. For illustrative purposes, we estimate and briefly discuss the classical regression estimation results, together with the pitfalls of this modelling approach in the present context.

Table 2.3 presents the estimation results of the intervention effects on the reported crime rate series based on single regression models with only a constant (level), only a trend (trend), and with trend and seasonal explanatory variables (trend + seasonal). We also include a dummy variable in December 1989 in burglary and larceny series, in order to control for an outlying observation in this year. Looking at the Table 2.2, this outlying observation might be due to the Virginia Firearms Transaction Programme, introduced in 1989.[18] For comparison purposes with later sections, the estimation results in Table 2.3 are produced by fitting deterministic state space models (with deterministic level, trend, and a seasonal component) to the analysed data. Deterministic state space models have its counterparts in classical linear regression (CLR) analysis, such that the deterministic local level model is equivalent to fitting a regression model to a constant, the deterministic local linear trend model is equivalent to fitting a regression model to a constant and a linear time trend, while the deterministic trend and seasonal model is equivalent to fitting a regression model to a constant, a linear time trend, and eleven dummies capturing the seasonal effects. For example, deterministic local level state model computes the mean and variance of the observed time series. The estimation results are generated by the numerical maximisation of the likelihood function, but no iterations are required for the estimations of parameters of the deterministic state space models. For more details on equivalence between the deterministic STS models and CLR models, see Commandeur and Koopman (2007). Structural time series models will be defined and discussed in Section 2.5.

In Table 2.3, the estimated coefficient δ for the intervention effect is reported together

[18]Virginia Firearms Transaction Programme is a criminal history records check to be conducted on persons purchasing a firearm from a licensed dealer. This programme provides a timely, point-of-sale check by allowing the dealer to contact the State Police via telephone or computer terminal. State Police provide the dealer with an approval or disapproval, usually within several minutes. Sales are disapproved if the background check reveals any state or federal prohibitor making the person legally ineligible to purchase a firearm. Virginia was the first state in the nation to establish such an "instant background check". Source: Crime in the Commonwealth 1988-1998.

with its t-test. Further, diagnostic tests for the assumptions of independence, homoscedasticity, and normality of residuals are presented. A more detailed definition, interpretation, and critical values for these diagnostic tests are given in the Appendix 2.A.1. Estimation results in Table 2.3 suggest significant intervention effects in almost all analysed crime series, with the exception of aggravated assault (coefficient significant only for the level model). We conclude that abolished parole and reformed sentencing in Virginia had negative effect on the recorded crime rates. We have drawn the same conclusion by looking at the graphs in Section 2.3.3. However, a key assumption of the classical regression analysis is not considered here, because the residuals in the fitted models are not randomly distributed. Presented diagnostic test statistics are not satisfactory. For example, values for the Q statistics are large, indicating that the residuals in the fitted models are serially correlated. First order residual autocorrelation, $r(1)$, is positive and significantly different from zero in all estimated models, which means that a positive residual tends to be followed by one or more other positive residuals, and a negative residual tends to be followed by one or more other negative residuals. In this case, the error variance for standard statistical test is seriously underestimated (Jovičić, 1989; Johnston and DiNardo, 1997; Verbeek, 2000), which leads in turn to a large overestimation of the F- and t-tests. Large t-test statistic brings about overly optimistic conclusions about the linear relationship between the analysed crime rates and intervention variable. Other diagnostics are also not satisfactory, indicating that the classical regression analysis is not appropriate for seasonal time series such as crime rates. The size of the estimated intervention effects is not reliable, although we have obtained a rough idea of the linear trend in the series. In the next part, we shall deal with the serial correlation in the error term, in order to get a better fitted model.

2.4.2 Regression models with ARMA errors

Since time series are by nature subject to serial correlation, the standard errors of the OLS parameter estimates are *biased*. As a result of this bias, t-test used to test the null hypothesis may overstate the statistical significance of an impact. For this reason, the time series should not be analysed by means of ordinary least squares regression methods. On the assumption that the time series corrected for *fixed* trend and seasonal effects is stationary, we may consider autoregressive-moving average (ARMA) processes for the explicit modelling of the serial dependence in the error term. The generalised regression model with disturbances, now denoted u_t rather than ε_t, that follow a stationary ARMA process is given by

$$y_t = x_t'\beta + \delta B_t + u_t, \tag{2.5}$$

where u_t is modelled by the ARMA model that can be represented as

$$u_t = \phi_1 u_{t-1} + \ldots + \phi_p u_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}, \tag{2.6}$$

for fixed integers p and q. The disturbance u_t is a white noise term (serially uncorrelated across time) and will be assumed normally distributed. It is important to note that setting an AR model by adding lagged dependent variables to the right hand side of the

Table 2.3: Estimated interventions for regression models

		interv	t-test	irr89:12	t-test	$Q(15)$	$r(1)$	$H(91)$	N	loglik	p.e.v.	R_s^2	AIC
BURGLAR	level	-25.77	-28.50			994.68	0.77	1.32	17.85	-557.15	55.94	-4.51	4.04
	trend	-7.94	-6.07			431.24	0.63	0.80	1.36	-473.77	29.35	-1.89	3.40
	trend + seas + irr	-7.48	-7.06	-10.24	-2.27	799.56	0.75	1.59	2.57	-416.30	18.32	-0.80	3.02
LARCEN	level	-45.71	-13.06			1410.70	0.83	1.05	4.45	-927.94	837.78	-8.42	6.75
	trend	-26.22	-3.82			1063.40	0.78	3.62	0.19	-926.27	807.65	-8.08	6.71
	trend + seas + irr	-24.43	-4.21	-12.29	-0.50	2321.5	0.89	13.11	27.98	-860.24	549.98	-5.19	6.42
MVTHEF	level	-3.71	-7.06			1642.50	0.84	0.35	11.39	-408.80	18.94	-7.66	2.95
	level + seas	-3.71	-7.72			2483.40	0.91	0.37	15.03	-387.43	15.19	-5.94	2.81
	trend	-3.82	-3.63			968.87	0.77	3.58	4.17	-414.31	18.98	-7.68	2.97
	trend + seas	-3.55	-3.69			1840.60	0.85	9.11	3.91	-393.02	15.22	-5.96	2.82
ROBBER	level	-1.53	-7.09			1143.90	0.77	0.82	12.29	-164.22	3.18	-2.79	1.17
	level + seas	-1.53	-8.03			1785.20	0.82	1.19	7.31	-143.71	2.38	-1.84	0.96
	trend	-2.03	-4.74			963.28	0.75	2.13	1.54	-169.72	3.16	-2.77	1.17
	trend + seas	-1.89	-4.96			1557.40	0.80	3.83	4.64	-149.64	2.37	-1.83	0.97
AGRASL	level	-0.92	-3.64			971.00	0.74	0.75	5.08	-209.35	4.42	-2.62	1.50
	level + seas	-0.92	-4.61			1784.10	0.76	0.99	7.96	-157.19	2.64	-1.16	1.06
	trend	-0.71	-1.41			846.31	0.71	1.70	4.14	-215.48	4.42	-2.62	1.50
	trend + seas	-0.64	-1.59			1488.50	0.71	3.61	0.22	-163.33	2.64	-1.16	1.07
MURDER	level	-0.19	-12.59			300.20	0.35	0.71	7.50	570.96	0.01	0.15	-4.20
	level + seas	-0.19	-12.77			342.68	0.36	0.76	8.40	533.08	0.01	0.21	-4.19
	trend	-0.17	-5.72			228.99	0.31	1.16	6.86	562.09	0.01	0.15	-4.19
	trend + seas	-0.16	-5.65			266.95	0.32	1.28	7.89	524.33	0.01	0.21	-4.18
RAPE	level	-0.40	-9.02			423.46	0.49	0.53	2.78	268.01	0.14	-0.52	-1.98
	level + seas	-0.40	-12.10			387.24	0.37	0.73	8.71	315.84	0.07	0.19	-2.53
	trend	-0.44	-4.90			398.94	0.47	0.77	1.79	260.11	0.14	-0.52	-1.97
	trend + seas	-0.44	-6.58			284.56	0.33	1.42	3.39	307.75	0.07	0.19	-2.53

Note: Sample size is 276; $Q(p)$ is portmanteau Box-Ljung test statistic; N is Bowman-Shenton normality test; $r(1)$ is the value of autocorrelation at lag 1; H-statistic tests for homoscedasticity of residuals; log-likelihood, one-step ahead prediction error variance (p.e.v.), R_s^2 value is corrected for seasonal means, and Akaike Information Criterion (AIC) are goodness-of-fit criteria.

equation (2.5) confounds the dynamics of the series with those of the intervention effect. For example,

$$y_t = \phi y_{t-1} + x_t'\beta + \delta B_t + \zeta_t, \qquad \zeta_t \sim N(0, \sigma_\zeta^2), \qquad t = 1, \ldots, n, \tag{2.7}$$

implies that there is a geometric distributed lag in the effect of the intervention. If we want to avoid this, the intervention variable has to be defined as $B_t - \phi B_{t-1}$, but this is exactly the same as letting the disturbance in (2.5) follow a first-order AR process and it requires non-linear estimation with respect to ϕ.

The choice of p and q is usually done on a trial and error basis, following the three-stage iterative procedure of *model selection* as defined by Box and Jenkins (1976). In the first stage, *identification*, a preliminary model is selected on the base of the appearance of the autocorrelation function (ACF) and the partial autocorrelation function (PACF).[19] Given p and q, the parameters are then estimated by the maximum likelihood (ML) method. As a result of the *estimation* stage, residuals are computed and used as the basis for *diagnostic checking*. A direct plot of the residuals can be a preliminary test for departures from randomness. A statistically valid check is based on testing if sample autocorrelations are not statistically different from zero. To check this, portmanteau Ljung-Box-Pierce $Q\,(p)$ test statistic is used (Harvey, 1993) and tested against an appropriate significance value from the χ^2_{P-p-q} distribution.[20] A significant departure from randomness indicates that the model is inadequate. In this case, a return to the identification stage is necessary and the complete three-stage process is repeated until a suitable formulation of the model is found. The model with the best fit and with satisfactory diagnostic statistics for normality, serial independence and homoskedasticity is chosen as the optimal model.

Table 2.4 describes the results of regression models with *fixed* level, trend, and seasonal explanatory variables. Burglary and larceny series are modelled with constant level, trend, and seasonal components, whereas motor vehicle theft, robbery, aggravated assault, murder, and rape series are modelled with constant level and seasonal components. Figures 2.7 and 2.8 show that fitting a linear trend through these series produced almost flat lines, which only differ in levels before and after the new legislation. In order to control for the serial correlation in the error term, we let the disturbances follow a first- or a second-order AR process. We also include a dummy variable in December 1989 in burglary and larceny series, in order to control for an outlying observation in this year. Looking at the Table 2.2, this outlying observation might be due to the Virginia Firearms Transaction Programme, introduced in 1989.

Estimation results in Table 2.4 do not suffer from serial correlation in the error term, given the low values for Q. It can also be seen that the estimations results and diagnostics do not differ significantly if we let the disturbances follow a first- or a second-order AR process.[21] In comparison to the Table 2.3, estimated effects of new legislation are not all statistically significant. Negative effects of the new legislation are found for burglary,

[19] Identification is based on the ACF and PACF which are estimated from the pre-intervention series. The pre-intervention series is examined, because the intervention can change the form of the time series.

[20] P should be chosen in such a way so as to be reasonably large compared to $p + q$. A rule of thumb is to set P equal to $\sqrt{n}$.

[21] Whether we let errors follow AR1 or AR2 process for larceny series plus dummy variable for an outlying observation in December 1989, does not change the results significantly.

larceny, robbery, murder, and rape. Significant negative results of the new legislation are only found for burglary and murder series. Positive effects of the new legislation are found for motor vehicle theft and aggravated assault, but these estimation results are not significant. It may therefore be concluded that regression results with weak diagnostic statistics can easily produce spurious results. However, the classical regression model has provided a rough idea of the linear trend in the series. In order to assess to what extent these preliminary results are reliable, we conduct a more rigorous time series study based on two model-based methodologies: the Box-Jenkins ARIMA analysis and the structural time series analysis.

2.4.3 Box-Jenkins ARIMA model approach

Most economic time series encountered in practice are non-stationary. The ARIMA methodology, developed by Box and Jenkins (1976), is based on the idea that series can be made stationary by operations such as differencing. In the case that the time series is difference stationary, it can be transformed to stationarity by taking first or second differences (or when needed, higher order differences), possibly in combination with seasonal differencing. Formally we define first and second differences by

$$\Delta y_t = y_t - y_{t-1}, \qquad \Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - 2y_{t-1} + y_{t-2},$$

respectively, and seasonal differences by

$$\Delta_s y_t = y_t - y_{t-s},$$

where s is the seasonal length (for example, $s = 12$ for a monthly time series). After taking the appropriate differences, the resulting time series will exhibit features of a stationary time series so that the appropriate ARMA(p, q) process can be used to model the remaining serial correlation in the series.

There are different ways in which one can decide on the appropriate degree of differencing (Harvey, 1993). The preliminary approach is to plot the data in levels and differences and to form an opinion according to which plots show trending movements. In the next step, an examination of the correlogram is necessary. For a stationary process, the main feature of the correlogram is that the autocorrelations tend toward zero as the lag increases. Finally, the unit root tests provide a more formal approach to determining the degree of differencing. More on unit root testing procedure can be found in Hamilton (1994).

Once the observations have been differenced, estimation and diagnostic checking my proceed as for an ARMA model. The proposed model

$$(1 - \phi_1 B - \ldots - \phi_p B^p) D(B) u_t = (1 + \theta_1 B + \ldots + \theta_q B^q) \varepsilon_t. \tag{2.8}$$

is called an *autoregressive-integrated-moving average process* of order (p, d, q) and denoted as ARIMA(p, d, q). In the above expression, B is the back shift or lag operator such that $B^k u_t = u_{t-k}$ and p and q are fixed integers. The difference function in the back shift operator $D(B)$ contains unit root coefficients, that is $\Delta^d = (1 - B)^d$. The disturbance

Table 2.4: Estimated interventions for a regression plus AR error models

		interv	t-test	irr89:12	t-test	$Q(12)$	$r(1)$	$H(87)$	N	loglik	p.e.v.	R_s^2	AIC
	trseas + ar1	-5.90	-2.40			42.30	0.04	0.75	15.19	-322.77	8.49	0.20	2.25
BURGLAR	trseas + ar2	-5.90	-2.40			42.32	0.04	0.75	15.18	-322.78	8.49	0.20	2.25
	trseas + ar2 + irr	-5.78	-2.41	-10.22	-4.08	55.21	0.04	1.02	0.50	-314.50	7.99	0.25	2.19
	trseas + ar1	-10.19	-1.32			30.00	0.01	1.17	40.92	-600.35	71.65	0.23	4.38
LARCEN	trseas + ar2	-10.02	-1.50			65.00	0.21	1.39	21.75	-605.94	74.67	0.20	4.42
	trseas + ar2 + irr	-9.40	-1.24	-31.50	-4.22	36.84	0.06	1.69	6.83	-590.09	67.08	0.29	4.32
	llseas + ar1	1.07	0.83			23.11	0.06	0.67	0.64	-123.65	1.95	0.15	0.77
MVTHEF	llseas + ar2	1.12	0.87			23.41	0.06	0.67	0.65	-123.58	1.94	0.15	0.77
	llseas + ar1	-0.27	-0.42			33.89	0.05	1.01	1.71	10.75	0.70	0.20	-0.26
ROBBER	llseas + ar2	-0.06	-0.09			35.90	0.05	1.02	1.97	10.75	0.70	0.20	-0.26
	llseas + ar1	0.05	0.07			22.51	0.05	0.97	4.04	-13.07	0.84	0.34	-0.07
AGRASL	llseas + ar2	0.18	0.30			22.89	0.11	0.98	4.67	-13.80	0.84	0.34	-0.07
	llseas + ar1	-0.12	-2.39			12.72	0.04	0.61	5.04	564.16	0.01	0.43	-4.46
MURDER	llseas + ar2	-0.11	-2.32			15.08	0.07	0.64	5.74	564.36	0.01	0.43	-4.47
	llseas + ar1	-0.18	-1.55			4.27	0.00	0.64	5.10	353.13	0.05	0.44	-2.86
RAPE	llseas + ar2	-0.13	-1.11			3.24	0.02	0.65	4.12	353.71	0.05	0.45	-2.86

Note: Sample size is 276; trseas denotes a specification with constant level, trend, and seasonal components; llseas denotes a specification with constant level and seasonal components; ar1 and ar2 denote first- and second-order autoregressive processes, respectively; $Q(p)$ is portmanteau Box-Ljung test statistic; N is Bowman-Shenton normality test; $r(1)$ is the value of autocorrelation at lag 1; H-statistic tests for homoscedasticity of residuals; log-likelihood, one-step ahead prediction error variance (p.e.v.), R_s^2 value is corrected for seasonal means, and Akaike Information Criterion (AIC) are goodness-of-fit criteria.

ε_t is a white noise term (serially uncorrelated across time) and will be assumed normally distributed.

The resulting impact assessment model can be represented as

$$y_t = \delta B_t + u_t, \tag{2.9}$$

where B_t is the step intervention and u_t is modelled as the ARIMA(p, d, q) process. The methodology for analysing time series intervention effects in the context of ARIMA models can be found in Box and Tiao (1975).

An extension of the ARIMA(p, d, q) model (2.8) is required to analyse time series with an evolving seasonal process. This can be handled by the use of seasonal differences. The first step is to generalise (2.8) to:

$$\Phi\left(B^s\right)\phi\left(B\right)\Delta^d\Delta_s^D u_t = \Theta\left(B^s\right)\theta\left(B\right)\varepsilon_t, \tag{2.10}$$

for $t = d+sD+1, \ldots, n$, where D is the degree of seasonal differencing; we have $\Delta = 1-B$ and $\Delta_s = 1 - B^s$. The lag polynomial functions are given by

$$\phi\left(B\right) = \left(1 - \phi_1 B - \ldots - \phi_p B^p\right), \qquad \Phi\left(B^s\right) = 1 - \Phi_1 B^s - \cdots - \Phi_P B^{P_s}, \tag{2.11}$$

$$\theta\left(B\right) = \left(1 + \theta_1 B + \ldots + \theta_q B^q\right), \qquad \Theta\left(B^s\right) = 1 + \Theta_1 B^s + \cdots + \Theta_Q B^{Q_s}. \tag{2.12}$$

This model is known as a *multiplicative seasonal* ARIMA model of order $(p, d, q) \times (P, D, Q)_s$.

The most popular model within the class of seasonal ARIMA models has become known as the 'Airline model' since it was originally fitted to a monthly series on UK airline passenger totals. The model is of order $(0, 1, 1) \times (0, 1, 1)_s$ with no constant and it is written as

$$(1 - B)(1 - B^s)u_t = (1 - \theta B)(1 - \Theta B^s)\varepsilon_t, \tag{2.13}$$

where s is the seasonal length ($s = 4$ for quarterly data and $s = 12$ for monthly data). The model has been found extremely useful in practice, because it has only a few parameters to estimate and fits data with pronounced seasonal effect generally well. In our empirical approach, airline model estimation results can serve well for comparison purposes. As it will be shown later in Section 2.5.3, the basic structural model has a reduced form which is quite close to the airline model.

Table 2.5 describes the results of the ARIMA models, based on the airline model specification. In addition, we report the results of the ARIMA $(0, 1, 1) \times (0, 1, 1)_{12}$ model. The choice of the other seasonal ARIMA model is based on the examination of the appropriate correlograms and parsimonicity. Similar to the regression results with ARMA errors described in Table 2.4, a statistically significant negative effect of the legislation is found for burglary, murder, and rape. However, when the airline model is considered, no significant intervention effect is detected. When we fit the airline model to the analysed series, positive effects of the legislation are found for motor vehicle theft, robbery, and aggravated assault, but these estimates are not statistically significant. The results in Table 2.5 do illustrate the sensitivity of the results with respect to the choice of differencing parameters d and D as part of a Box-Jenkins ARIMA analysis.

Table 2.5: Estimated interventions for ARIMA models

		interv	z-test	irr89:12	t-test	$Q(12)$	$r(1)$	loglik	AIC
BURGLAR	ARIMA(1,0,0)(0,1,1)$_{12}$	-6.57	-2.44	-13.12	-6.68	34.08	-0.27	-666.65	1361.19
	Airline model	-3.69	-0.89	-12.64	-6.06	14.24	0.04	-648.16	1306.31
LARCEN	ARIMA(1,0,0)(0,1,1)$_{12}$	-11.42	-1.40	-40.59	-11.37	37.85	-0.31	-949.80	1909.61
	Airline model	-6.31	-0.77	-37.52	-7.53	12.76	0.00	-928.30	1866.60
MVTHEF	ARIMA(1,0,0)(0,1,1)$_{12}$	-0.66	-0.30			32.74	-0.23	-496.87	1001.74
	Airline model	1.91	1.59			19.17	0.08	-481.22	970.44
ROBBER	ARIMA(1,0,0)(0,1,1)$_{12}$	-1.09	-1.44			32.65	-0.23	-349.46	706.92
	Airline model	0.61	1.17			10.81	0.05	-332.61	673.22
AGRASL	ARIMA(1,0,0)(0,1,1)$_{12}$	-0.64	-0.89			53.07	-0.20	-394.38	796.76
	Airline model	0.19	0.15			8.37	0.05	-367.51	743.02
MURDER	ARIMA(1,0,0)(0,1,1)$_{12}$	-0.16	-4.77			52.02	-0.09	186.24	-364.49
	Airline model	-0.09	-1.17			11.62	0.03	201.13	-394.27
RAPE	ARIMA(1,0,0)(0,1,1)$_{12}$	-0.38	-4.73			51.06	-0.07	-29.23	66.46
	Airline model	-0.13	-0.89			4.62	-0.01	-9.68	27.35

Note: Sample size, ARIMA(1,0,0)(0,1,1)$_{12}$ = 264; Airline model = 263; $Q(p)$ is portmanteau Box-Ljung test statistic; N is Bowman-Shenton normality test; $r(1)$ is the value of autocorrelation at lag 1; H-statistic tests for homoscedasticity of residuals; log-likelihood, one-step ahead prediction error variance (p.e.v.), R_s^2 value is corrected for seasonal means, and Akaike Information Criterion (AIC) are goodness-of-fit criteria.

2.5 The structural time series approach

Structural time series models are formulated directly in terms of components of interest, that is, trend, seasonal, and error components, plus other relevant terms, which have a direct interpretation and allow for time-varying parameters. In the reduced form, STS models are ARIMA models with parameter restrictions, hence the results can also have ARIMA interpretation. Major advantage of the STS methods is that sources of shocks are more easily identified by structural modelling than by the ARIMA models. Namely, "ARIMA models tend to smear departures in one component into diagnostics for several aspects of the model," Atkinson et al. (1997). The STS models are estimated both in the univariate and multivariate domains. Following the approach of Harvey (1996), in the multivariate domain we estimate multivariate STS models with control groups.

2.5.1 The basic structural time series model

The basic model for representing a time series is the additive model:

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \qquad t = 1, \ldots, n, \tag{2.14}$$

also known as the Classical Decomposition or Basic Structural Model (BSM). In model (2.14), y_t is observation, μ_t is a slowly changing component (trend), γ_t is a periodic component (seasonal), and ε_t is irregular component (error). In a *Structural Time Series Model (STSM)* or *Unobserved Components Model (UCM)*, the right-hand side components are modelled explicitly as stochastic processes. If we consider a simple form of model (2.14) in which μ_t is a random walk, no seasonal is present and all random variables are normally distributed, then we obtain the *local level* (LL) or random walk plus noise model, as given by

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, \qquad & \varepsilon_t &\sim \mathcal{NID}\left(0, \sigma^2_\varepsilon\right), \\ \mu_{t+1} &= \mu_t + \xi_t, \qquad & \xi_t &\sim \mathcal{NID}\left(0, \sigma^2_\xi\right), \end{aligned} \tag{2.15}$$

for $t = 1, \ldots, n$, where the ε_t's and ξ_t's are mutually independent and are independent of α_1. In this model the *level* of the estimated "true" development is allowed to vary over time, i.e., the *level* is only fixed *locally*. Hence the name of the model. The local level model is a simple example of a *linear Gaussian state space model*. Since all components are assumed to be stochastic, the measurement errors are included in both equations. The variable μ_t is called the *state* and is unobserved, together with error terms ε_t and ξ_t. The object of the methodology is to study the development of the state over time using the observed values $y_1, \ldots, y_n$. Hence, the first equation is called the *observation equation*. The model is incomplete without a specification for $\mu_1 \sim \mathcal{N}(a, P)$. Parameters to be estimated are: $a, P, \sigma^2_\varepsilon, \sigma^2_\eta$. The model is estimated using the state space form and Kalman filter, as described in Harvey (1989).

Taking first differences in (2.15) yields a process in which the first order autocorrelation is $-1\,/\,(2+q)$, where $q = \sigma^2_\xi / \sigma^2_\varepsilon$ is *signal-to-noise ratio*, while the remaining higher order autocorrelations are all zero. Hence the autocorrelation function (ACF) is the same as that of an MA(1) process and so the local level model has an ARIMA(0,1,1) reduced form.

The local level model is a simple form of a structural time series model. By adding a slope term ν_t, which is generated by a random walk, we can derive the *local linear trend* (LLT) model:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \mathcal{NID}\left(0, \sigma_\varepsilon^2\right), \\ \mu_{t+1} &= \mu_t + \nu_t + \xi_t, & \xi_t &\sim \mathcal{NID}\left(0, \sigma_\xi^2\right), \\ \nu_{t+1} &= \nu_t + \zeta_t, & \zeta_t &\sim \mathcal{NID}\left(0, \sigma_\zeta^2\right). \end{aligned} \tag{2.16}$$

The local linear trend model contains two state equations: one for modelling the level, and one for modelling the slope. If $\xi_t = \zeta_t = 0$, then $\nu_{t+1} = \nu_t = \nu$, and $\mu_{t+1} = \mu_t + \nu$ so that the trend is exactly linear and (2.16) reduces to the deterministic linear trend plus noise model. The form (2.16) with $\sigma_\xi^2 > 0$ and $\sigma_\zeta^2 > 0$ allows the trend level and slope to vary over time. Special cases of STS models and their ARIMA reduced forms are presented in Table (2.6).

Table 2.6: Special Cases of the STS Models

	σ_ε^2	σ_η^2	σ_ξ^2	Model	Reduced form
LL	>0	>0		stochastic level	ARIMA(0,1,1)
	>0	0		$y_t = \mu_1 + \varepsilon_t$ (CLRM)	
	0	>0		$y_{t+1} = y_t + \eta_t$ (pure RW)	
LLT	>0	>0	>0	stochastic level & slope	ARIMA(0,2,2)
	>0	0	0	deterministic linear trend	
	>0	>0	0	trend is a RW with constant drift	
	>0	0	>0	trend is a smooth curve or IRW	

CLRM = classical linear regression model; RW = random walk;
IRW = integrated random walk

In the structural time series methodology, a seasonal component can be modelled by adding it either to the local level model or to the local linear trend model. Just as we allowed a trend to be stochastic, the seasonal component can also be allowed to change over time. Various specifications for the seasonal component γ_t exist. For example, we can opt for a 'dummy variable' or 'trigonometric' form of stochastic seasonality. For our empirical analysis we adopt a trigonometric specifications since the statistical properties imply a smooth seasonal process and its parametrisation is flexible. Defining γ_{jt} as the effect of season j at time t, the stochastic specification of the trigonometric seasonal component is given by

$$\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{jt}, \tag{2.17}$$

where $[a]$ is the largest integer smaller than a and

$$\begin{aligned} \gamma_{j,t+1} &= \gamma_{jt} \cos \lambda_j + \gamma_{jt}^* \sin \lambda_j + \omega_{jt}, \\ \gamma_{j,t+1}^* &= -\gamma_{jt} \sin \lambda_j + \gamma_{jt}^* \cos \lambda_j + \omega_{jt}^*, \end{aligned}$$

with frequency $\lambda_j = 2\pi j/s$, for $j = 1, \ldots, [s/2]$. For monthly data, we have $s = 12$ and $j = 1, \ldots, 6$. The disturbances are serially and mutually uncorrelated and they are normally distributed with mean zero and variance matrix

$$Var\begin{pmatrix} \omega_{j,t} \\ \omega^*_{j,t} \end{pmatrix} = \begin{bmatrix} \sigma^2_\omega & 0 \\ 0 & \sigma^2_\omega \end{bmatrix}.$$

The extent to which the seasonal pattern change over time depends on a variance σ^2_ω. When this is zero, the trigonometric model will give the same deterministic pattern as the dummy variable seasonal model. However, the trigonometric seasonal variable model allows seasonal pattern to evolve more smoothly than the dummy variable seasonal model. The ω_{jt}'s and ω^*_{jt}'s are restricted to have a common variance σ^2_ω, but this restriction can be relaxed so that they have different variances for $j = 1, \ldots, [s/2]$.

When we combine seasonal effects, together with trend and irregular components, we get the basic structural model in (2.14), where the stochastic trend component, μ_t, is defined as in (2.16). The BSM has a reduced form which is quite close to the airline model of Box and Jenkins, namely the seasonal ARIMA of order $(0, 1, 1,) \times (0, 1, 1)_s$. The special case of BSM when the seasonal and slope disturbance variances are equal to zero is equivalent to an airline model with the seasonal moving average parameter, Θ, equal to minus one.

Intervention effects can be incorporated in the structural time series model framework. To account for the change in level due to an intervention at time t, we add the intervention regression effect to model (2.14) and we obtain

$$y_t = \mu_t + \gamma_t + \delta B_t + \varepsilon_t, \qquad t = 1, ..., n. \tag{2.18}$$

where the intervention effect B_t is defined in (2.2) and δ is a fixed but unknown coefficient. It follows that δ measures the change in the level of the series at a known time t.

Explanatory variables can also be incorporated in the structural time series modelling framework. Suppose we have k regressors $x_{1t}, \ldots, x_{kt}$ with unknown regression coefficients $\beta_1, \ldots, \beta_k$ which are constant over time. By adding the regression effects into model 2.18 we obtain

$$y_t = \mu_t + \gamma_t + \sum_{j=1}^{k} \beta_{jt} x_{jt} + \delta B_t + \varepsilon_t, \qquad t = 1, ..., n. \tag{2.19}$$

where the β_{jt}'s are unknown coefficients. The resulting model can be put into the state space form and estimated by Kalman filter. The fully specified local level model that we are going to estimate is shown in Table 2.7, where coefficient on intervention is treated as fixed regression parameter.

Summarizing, the STS models gain in flexibility as compared to other time series models because the stochastic formulation allows mean, trend, and seasonality to evolve over time. Further, different specifications for the level, the trend, and the seasonal can be freely combined. Other components of interest, like cycles, explanatory variables, interventions effects, outliers, are easily added to the basic structural model. Mix of

Table 2.7: Fully Specified LL Model

	Equation	Error term
observation equation	$y_t = \mu_t + \gamma_t + \beta_t x_t + \delta_t I_t + \varepsilon_t$	$\varepsilon_t \sim \mathcal{NID}(0, \sigma_\varepsilon^2)$
level	$\mu_{t+1} = \mu_t + \xi_t$	$\xi_t \sim \mathcal{NID}(0, \sigma_\xi^2)$
seasonal	$\gamma_{j,t+1} = \gamma_{jt} \cos \lambda_j + \gamma_{jt}^* \sin \lambda_j + \omega_{jt}$	$\omega_{jt} \sim \mathcal{NID}(0, \sigma_\omega^2)$
seasonal	$\gamma_{j,t+1}^* = -\gamma_{jt} \sin \lambda_j + \gamma_{jt}^* \cos \lambda_j + \omega_{jt}^*$	$\omega_{jt}^* \sim \mathcal{NID}(0, \sigma_\omega^2)$
expl. var. coeff.	$\beta_{t+1} = \beta_t + \tau_t$	$\tau_t \sim \mathcal{NID}(0, \sigma_\tau^2)$
intervention coeff.	$\delta_{t+1} = \delta_t + \rho_t$	$\rho_t \sim \mathcal{NID}(0, \sigma_\rho^2)$
	$\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{jt}, \;\; \lambda_j = 2\pi j/s, \;\; j = 1, \ldots, [s/2]$	
	$\beta_{t+1} = \beta_t = \beta, \;\; \sigma_\tau^2 = 0$	
	$\delta_{t+1} = \delta_t = \delta, \;\; \sigma_\delta^2 = 0$	

quarterly and monthly series is easily handled (problem of missing values). As it will be shown later, we can straightforwardly generalise the analysis of one time series to the simultaneous analysis of two or more time series (multivariate analysis). In the reduced form, STS models are ARIMA models with parameter restrictions (often, these restrictions are very useful). Unlike in ARIMA analysis, we model non-stationarity directly.

In words of Harvey (1996), "the basic idea of structural time series models is that they are set up as regression models in which the explanatory variables are functions of time, but with coefficients that change over time. Model specification proceeds on the basis that the researcher generally has a good idea of what components to include from the outset, though any model will always be subject to diagnostic checking in much the same way as an ARIMA model." More details about the structural time series models and its dynamic properties are given in, for example, Harvey (1989) and Durbin and Koopman (2001).

2.5.2 State space representation, filtering and smoothing

All univariate state space models can be expressed using matrix algebra in the following general format:

$$y_t = z_t' \alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, \sigma_\varepsilon^2) \tag{2.20}$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \qquad \eta_t \sim \mathcal{NID}(0, Q_t) \tag{2.21}$$

for $t = 1, \ldots, n$. The terms y_t and ε_t are scalars as before. However, the remaining terms in (2.20) and (2.21) denote vectors and matrices. In particular, z_t is an $m \times 1$ *observation* or *design* vector, T_t is an $m \times m$ *transition* matrix, α_t is an $m \times 1$ *state* vector, and m denotes the number of elements in the state vector. In many state space models R_t in (2.21) is simply the identity matrix of order $m \times m$. However, in various models it is of order $m \times r$ with $r < m$, and consists of the first r columns of the identity matrix I_m. In this case R_t is called a *selection* matrix since it selects the rows of the state equation which have non-zero disturbance terms. Finally, the $r \times 1$ vector η_t contains the r state disturbances with zero means, and unknown variances collected in an $r \times r$ diagonal matrix Q_t. In this general formulation, equation (2.20) is called the *observation* of *measurement* equation, while equation (2.21) is called the *transition* or *state* equation.

By appropriate definitions of the vectors z_t, α_t, and η_t, and of the matrices T_t, R_t, and Q_t, all the STS models discussed in Section 2.5.1 can be derived as special cases of (2.20) and (2.21). In this section, the state space representation of the models discussed until now is presented.

The local level model is the simplest special case of (2.20) and (2.21). Since the state vector of the local level model consists of only one element (i.e., the level), $m = 1$ in this case. We can define

$$\alpha_t = \mu_t, \qquad \eta_t = \xi_t, \qquad z_t = T_t = R_t = 1, \qquad Q_t = \sigma^2_\xi,$$

(all of order 1×1) for $t = 1, \ldots, n$, it is easily verified that (2.20) and (2.21) simplifies into the local level model in (2.15).

The local linear trend model requires 2×1 state vector: one element for the level μ_t and one element for the slope ν_t. The matrix algebra of the local linear trend model in (2.16) looks as follows

$$\alpha_t = \begin{pmatrix} \mu_t \\ \nu_t \end{pmatrix}, \qquad \eta_t = \begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix}, \qquad T_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad z_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$Q_t = \begin{bmatrix} \sigma^2_\xi & 0 \\ 0 & \sigma^2_\zeta \end{bmatrix}, \qquad \text{and} \qquad R_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The local level model can also be extended with a stochastic trigonometric seasonal effect. The state vector α_t has the dimension $(s+1) \times 1$ and the complete state space form looks as follows

$$\alpha_t = \begin{pmatrix} \mu_t \\ \gamma_{1,t} \\ \gamma^*_{1,t} \\ \gamma_{2,t} \\ \gamma^*_{2,t} \\ \vdots \\ \gamma_{6,t} \end{pmatrix}, \qquad \eta_t = \begin{pmatrix} \xi_t \\ \omega_{1t} \\ \omega^*_{1t} \\ \omega_{2t} \\ \omega^*_{2t} \\ \vdots \\ \omega_{6t} \end{pmatrix},$$

$$T_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & cos\lambda_1 & sin\lambda_1 & 0 & 0 & & 0 & 0 & 0 \\ 0 & -sin\lambda_1 & cos\lambda_1 & 0 & 0 & & 0 & 0 & 0 \\ 0 & 0 & 0 & cos\lambda_2 & sin\lambda_2 & & 0 & 0 & 0 \\ 0 & 0 & 0 & -sin\lambda_2 & cos\lambda_2 & & 0 & 0 & 0 \\ \vdots & & & & & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & cos\lambda_{s^*} & sin\lambda_{s^*} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -sin\lambda_{s^*} & cos\lambda_{s^*} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

$$z_t = (1\ 1\ 0\ 1\ 0\ 1\ \dots\ 1\ 0\ 1)',$$

$$Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma_\omega^2 & 0 & 0 & 0 & & 0 \\ 0 & 0 & \sigma_\omega^2 & 0 & 0 & & 0 \\ 0 & 0 & 0 & \sigma_\omega^2 & 0 & & 0 \\ 0 & 0 & 0 & 0 & \sigma_\omega^2 & & 0 \\ \vdots & & & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_\omega^2 \end{bmatrix}, \text{ and } R_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 1 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 1 & & 0 \\ \vdots & & & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

where we assume that s is even, $s^* = \left[\frac{s-1}{2}\right]$, $\lambda_j = \frac{2\pi j}{s}$, $j = 1, \dots, [s/2]$. When s is odd, we have

$$z_t = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad T_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & cos\lambda_1 & sin\lambda_1 & 0 & 0 & & 0 & 0 \\ 0 & -sin\lambda_1 & cos\lambda_1 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & cos\lambda_2 & sin\lambda_2 & & 0 & 0 \\ 0 & 0 & 0 & -sin\lambda_2 & cos\lambda_2 & & 0 & 0 \\ \vdots & & & & & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & cos\lambda_{s^*} & sin\lambda_{s^*} \\ 0 & 0 & 0 & 0 & 0 & 0 & -sin\lambda_{s^*} & cos\lambda_{s^*} \end{bmatrix}.$$

The local level model extended with one explanatory variable looks in matrix algebraic formulation as follows

$$\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}, \qquad \eta_t = \xi_t, \qquad T_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad z_t = \begin{pmatrix} 1 \\ x_t \end{pmatrix},$$

$$Q_t = \sigma_\xi^2, \qquad \text{and} \qquad R_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Similarly, the local level model with an intervention variable has the following matrix representation

$$\alpha_t = \begin{pmatrix} \mu_t \\ \delta_t \end{pmatrix}, \qquad \eta_t = \xi_t, \qquad T_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad z_t = \begin{pmatrix} 1 \\ I_t \end{pmatrix},$$

$$Q_t = \sigma_\xi^2, \qquad \text{and} \qquad R_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Our fully specified local level and trigonometric seasonal model for monthly data ($s = 12$), together with a deterministic explanatory variable x_t and an intervention variable

I_t, shown in scalar notation in Table (2.7), can be put into the state space format in the following way

$$\alpha_t = \begin{pmatrix} \mu_t \\ \gamma_{1,t} \\ \gamma^*_{1,t} \\ \gamma_{2,t} \\ \gamma^*_{2,t} \\ \vdots \\ \beta_t \\ \delta_t \end{pmatrix}, \qquad \eta_t = \begin{pmatrix} \xi_t \\ \omega_{1t} \\ \omega^*_{1t} \\ \omega_{2t} \\ \omega^*_{2t} \\ \vdots \end{pmatrix},$$

$$T_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & cos\lambda_1 & sin\lambda_1 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 \\ 0 & -sin\lambda_1 & cos\lambda_1 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & cos\lambda_2 & sin\lambda_2 & & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -sin\lambda_2 & cos\lambda_2 & & 0 & 0 & 0 & 0 & 0 \\ \vdots & & & & & \ddots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & cos\lambda_{s^*} & sin\lambda_{s^*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -sin\lambda_{s^*} & cos\lambda_{s^*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$z_t = (1\ 1\ 0\ 1\ 0\ 1\ \dots\ 1\ 0\ 1\ x'_t\ B_t)',$$

$$Q_t = \begin{bmatrix} \sigma^2_\xi & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2_\omega & 0 & 0 & 0 & & 0 \\ 0 & 0 & \sigma^2_\omega & 0 & 0 & & 0 \\ 0 & 0 & 0 & \sigma^2_\omega & 0 & & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_\omega & & 0 \\ \vdots & & & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_\omega \end{bmatrix}, \quad \text{and} \quad R_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & & 0 \\ 0 & 0 & 1 & 0 & 0 & & 0 \\ 0 & 0 & 0 & 1 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 1 & & 0 \\ \vdots & & & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where we assume that s is even, $s^* = \left[\frac{s-1}{2}\right]$, $\lambda_j = \frac{2\pi j}{s}$, $j = 1, \dots, [s/2]$.

Once in the state space format, the state components can be estimated in different ways. The state estimates are considered for given values of hyperparameters (i.e., the variances of the irregular and of the state disturbances) and for given initial values of the state components. Following the exposition of Commandeur and Koopman (2007), the estimations of the state vector are carried out by performing two passes through the data:

i. a *forward* pass, from $t = 1, \dots, n$, using a recursive algorithm known as the *Kalman filter* that is applied to the observed time series;

ii. a *backward* pass from $t = n, \dots, 1$, using recursive algorithm known as *state and disturbance smoothers* that are applied to the output of the Kalman filter.

The main purpose of the Kalman filter is to update our knowledge of the system each time a new observation y_t is brought in. Hence, we want to obtain optimal values of the state at time point t, *only* considering past observations $\{y_1, y_2, \dots, y_{t-1}\}$. The backward pass through the data is only required for state and disturbance smoothing. in order to obtain estimated values of the state and disturbance vectors at time point t, considering *all* available observations $y_1, y_2, \dots, y_n$. All STSM estimation results that we present are smoothed estimates of the state vector for which all observations are used.

If we denote the set of past observation $\{y_1, y_2, \dots, y_{t-1}\}$ by Y_{t-1} and if we assume that the conditional distribution of α_t given Y_{t-1} is $N(a_t, P_t)$, given that a_t and P_t are known, the purpose of Kalman filter is to calculate a_{t+1} and P_{t+1} when y_t is brought in. The central formula in the recursive Kalman filter algorithm is:

$$a_{t+1} = a_t + K_t(y_t - z_t' a_t), \tag{2.22}$$

where a_t denotes the filtered state at time point t. For a full set of relations constituting Kalman filter, we refer the reader to Durbin and Koopman (2001).

For the local level model, equation (2.22) becomes $a_{t+1} = a_t + K_t(y_t - a_t)$. If we let $\upsilon_t = y_t - a_t$ for $t = 1, \dots, n$, the values of υ_t are called *one-step-ahead forecast (or prediction) errors*. They quantify the lack of accuracy of a_t in predicting the observed value of y_t ate time point t. The prediction errors are also denoted as *innovations* because they represent the new information in y_t that cannot be predicted from the past.

The value of K_t is equal to P_t/F_t where P_t denotes the filtered state estimation error variance, and F_t the variance of the one-step prediction errors υ_t. The K_t determines by *how much* the prediction error at time point t is allowed to influence the estimate of the state at time point $(t+1)$. For example, K_t equal to one allows the prediction error (innovation) to have a large impact on the next value of the state. On the other hand, K_t equal to zero does not allow that uncertainty of innovations have a large effect on the next value of the state. The value of K_t is therefore called the Kalman *gain* and can be interpreted as a compromise between the uncertainty of the state based on past observations $y_1, y_2, \dots, y_{t-1}$ and the uncertainty of the new observation y_t.

The prediction errors υ_t and their variances F_t are important for the maximisation of the log-likelihood function in state space methods. For univariate state space models the diffuse log-likelihood is defined as:

$$\log L_d = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=d+1}^{n}\left(\log F_t + \frac{\upsilon_t^2}{F_t}\right), \tag{2.23}$$

where d is the number of diffuse initial elements of the state. In state space methods the value of the log-likelihood function is maximised by simultaneously *minimising* the prediction errors υ_t and their variances F_t, whereas in classical linear regression models the estimates are obtained by minimising the *observation* errors ε_t and their variances σ_ε^2. Given the model specification, the aim is to obtain hyperparameter estimates that weight the past observations in an optimal way in order to provide the best prediction of the current observation.

2.5.3 Equivalence between STS and ARIMA models

Following the exposition of Commandeur and Koopman (2007), there are a number of important similarities between state space and ARIMA models. For example, consider the local level model as given by

$$y_t = \mu_t + \varepsilon_t, \tag{2.24}$$
$$\mu_t = \mu_{t-1} + u_t. \tag{2.25}$$

The first differences of y_t are equal to

$$\Delta y_t = y_t - y_{t-1} = \mu_t - \mu_{t-1} + \varepsilon_t - \varepsilon_{t-1}. \tag{2.26}$$

It follows from (2.25) that

$$\mu_t - \mu_{t-1} = u_t, \tag{2.27}$$

and substitution of (2.27) into (2.26) gives

$$\Delta y_t = y_t - y_{t-1} = u_t + \varepsilon_t - \varepsilon_{t-1}. \tag{2.28}$$

It can be shown that (2.28) is a stationary process which has the same correlogram as the MA(1) process. This implies that the local level model is equivalent to an ARIMA(0,1,1) model.

Next, consider the local linear trend model as given by

$$y_t = \mu_t + \varepsilon_t, \tag{2.29}$$
$$\mu_t = \mu_{t-1} + \nu_{t-1} + \xi_{t-1}, \tag{2.30}$$
$$\nu_t = \nu_{t-1} + \zeta_{t-1}. \tag{2.31}$$

Taking first differences of y_t in (2.29) gives

$$\Delta y_t = y_t - y_{t-1} = \mu_t - \mu_{t-1} + \varepsilon_t - \varepsilon_{t-1}, \tag{2.32}$$

and the second differences are therefore equal to

$$\begin{aligned}\Delta^2 y_t &= y_t - y_{t-1} - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2} \\ &= \mu_t + \varepsilon_t - 2(\mu_{t-1} + \varepsilon_{t-1}) + (\mu_{t-2} + \varepsilon_{t-2}) \\ &= (\mu_t - \mu_{t-1}) - (\mu_{t-1} - \mu_{t-2}) + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}. \end{aligned} \tag{2.33}$$

It follows from (2.30) that

$$\mu_t - \mu_{t-1} = \nu_{t-1} + \xi_{t-1}, \tag{2.34}$$

and

$$\mu_{t-1} - \mu_{t-2} = \nu_{t-2} + \xi_{t-2}, \tag{2.35}$$

Substitution of (2.34) and (2.35) in (2.33) gives

$$\begin{aligned}\Delta^2 y_t &= (\nu_{t-1} + \xi_{t-1}) - (\nu_{t-2} + \xi_{t-2}) + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2} \\ &= (\nu_{t-1} - \nu_{t-2}) + \xi_{t-1} - \xi_{t-2} + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon t - 2. \end{aligned} \tag{2.36}$$

Finally, it follows from (2.31) that

$$\nu_{t-1} - \nu_t = \zeta_{t-2}, \tag{2.37}$$

and substitution of (2.37) in (2.36) gives

$$\Delta^2 y_t = \zeta_{t-2} + \xi_{t-1} - \xi_{t-2} + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}. \tag{2.38}$$

It can be shown that (2.38) is a stationary process yielding the same correlogram as a MA(2) model. The local linear trend model is therefore equivalent to an ARIMA(0,2,2) model. A comprehensive overview of the equivalence between STS and ARIMA models is given in Harvey (1989) in Appendix 1. It should be also noted that ARIMA models can also be put into state space form and fitted by state space methods.

In spite of the close relationship between ARIMA and STS models, the two approaches to time series analysis are distinct. Major advantages of the STS methodology in compare to the ARIMA approach are: a) whereas trend and seasonal are explicitly modelled in the STS approach, in the ARIMA models they are removed from the series before any analysis is performed; b) in the ARIMA models the observed time series is differenced prior to the analysis, in order to obtain an approximation to stationary time series, while in the STS approach the time series is modelled directly in levels, whether stationary or not; c) mix of quarterly and monthly series (problem of missing values), stochastic explanatory variables, and multivariate model extensions are easily incorporated into the STS methodology. This is not so straightforward in a pure ARIMA context; d) in the reduced form, STS models are ARIMA models with parameter restrictions (these restrictions can be very useful). In words of Commandeur and Koopman (2007), "state space methods provide an explicit structural framework for the decomposition of time series in order to diagnose all the dynamics in the time series data simultaneously."

2.5.4 Structural time series analysis results

Residual analysis

From the one-step ahead prediction residual v_t and its variance F_t of the Kalman filter, we are able to construct the standardised residual $v_t \,/\, \sqrt{F_t}$. The series of residuals can be used to detect unusual movements in a fitted time series model. Atkinson (1985) suggests several methods of residual plots analysis in order to detect departures from the fitted model. He lists the following most useful plots:

i. A scatter plot of the response y against each of the explanatory variables x_j.

ii. A plot of the residuals against each explanatory variable in the model. The presence of a curved relationship suggests that a higher-order term, for example a quadratic in the explanatory variables, should be added to the model.

iii. A plot of the residuals against explanatory variables not in the model. A relationship would suggest that the variables should be included in the model.

iv. A plot of residuals against predicted values from the fitted model. If the variance of the residuals seems to increase with the predicted value, we should consider modelling, for example, a logarithmic transformation of the response y.

v. A normal plot of the residuals. After all the systematic variation has been removed from the data, the residuals should look pretty much like a sample from the normal distribution. A plot of the ordered residuals against the expected order statistics from a normal distribution provides a visual check of this assumption.

vi. Plots of the response y or residuals against time. These plots sometimes lead to the detection of unusual patterns due either to time itself or to other variables closely correlated with time.

Although it is very useful to investigate the prediction residuals, in time series analysis it is often not possible to detect outliers and structural breaks or distinguish between them on the basis of the prediction residuals. In the case of structural time series models, different residuals can be considered because the model contains of different equations with their own disturbances. The *auxiliary residuals* are smoothed estimates of the disturbances for particularly the level, slope and irregular components. Since they are smoothed estimates, auxiliary residuals are neither serially uncorrelated nor uncorrelated with each other. However, they can play a valuable role in separating information which are mixed up together in the one-step ahead prediction residuals. The level residual series enables the detection of a break in the time series while the irregular residuals can detect outliers.

Seasonality of the analysed crime series

Monthly crime data analysed in this chapter, like many other monthly time series data, are affected by short-term seasonal variations. Figure 2.11, showing trend and seasonal components of the property crime data, suggests that burglary had summer and Christmas peaks, but since 1995 end of the year peak is less pronounced. In 2006, the month of July had the greatest occurrences of burglary. Larceny and motor vehicle theft peak only in the summer months throughout the analysed period. Both offences have the highest frequency of occurrence in August. Robbery is the lowest in March and April, but peaks in the second half of the year. In 2006, the month of October had the greatest number of robberies reported. Figure 2.12, showing trend and seasonal components of the violent crime data, suggests that aggravated assault peaks in summer months, with July having the greatest number of offences reported. Rape is highest during spring and summer months, from April to August. Month of August has the highest occurrences of rape incidents. Murder is highest in the second half of the year (from June onward), with a dip in October. It peaks in June, July, August and September and again in November and December each year, where December has the greatest occurrences of murder. Hird and Ruparel (2007), analysing the seasonality of monthly recorded crime data in England and Wales, find that the violent assault offences and sexual offences peak in summer months and through in the winter months, whereas opposite is true for property crimes. In comparison to their findings, we find that burglary (after 1995), larceny, motor vehicle theft, aggravated assault and rape all peak in the summer months and through in the winter months. Opposite is true for robbery and murder, which peak in the winter (murder almost from June to December) and through in the summer months.

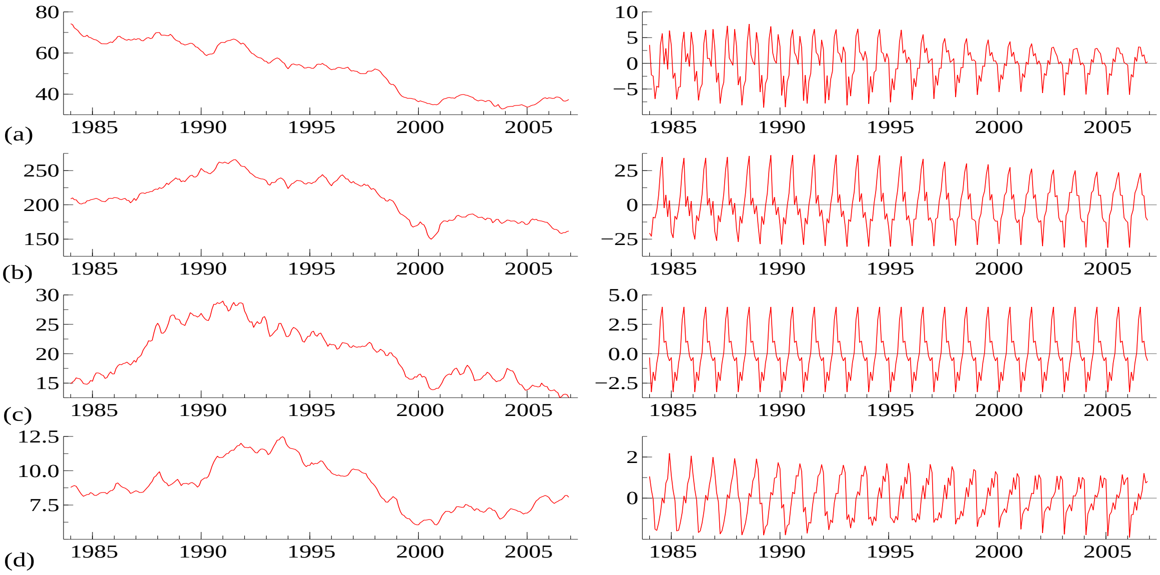

Figure 2.11: Trend and seasonal components in property crime rates: (a) burglary; (b) larceny; (c) motor vehicle theft; (d) robbery.

These findings can be interpreted from the perspective of economic models of criminality, developed by Becker (1968) and Ehrlich (1973), where criminals are seen as rational. The frequency of criminal events will tend to increase when the perceived gains from engagement in crime increase, ceteris paribus. This suggests that individual criminals might exhibit a considerable amount of mobility over time, as they seek those situations where perceived gains are greatest and/or the subjective probability of detection and arrest are the smallest. For example, rape incidents have greatest occurrences in summer months, when social interaction are at its highest level and climatic conditions make victims more available. On the other hand, en economic crime such as robbery peaks in the winter months, due to the increase in the cost of living and the facilitating environmental conditions during these months. Because of the variety of motives involved in murder, this crime is much less dependent on climatic conditions. For a discussion on the seasonality of violent crimes, see Landau and Fridman (1993).

Auxiliary residuals

In unobserved component models, it can also be useful to analyse the auxiliary residuals, that estimate the disturbances of each component. Auxiliary residuals are defined as the smoothed estimates of the disturbances associated with the components irregular, level, and slope; see the discussion in Section 2.5.4. Harvey and Koopman (1992) derive expressions for auxiliary residuals, and propose to use them to identify outliers; see also Durbin

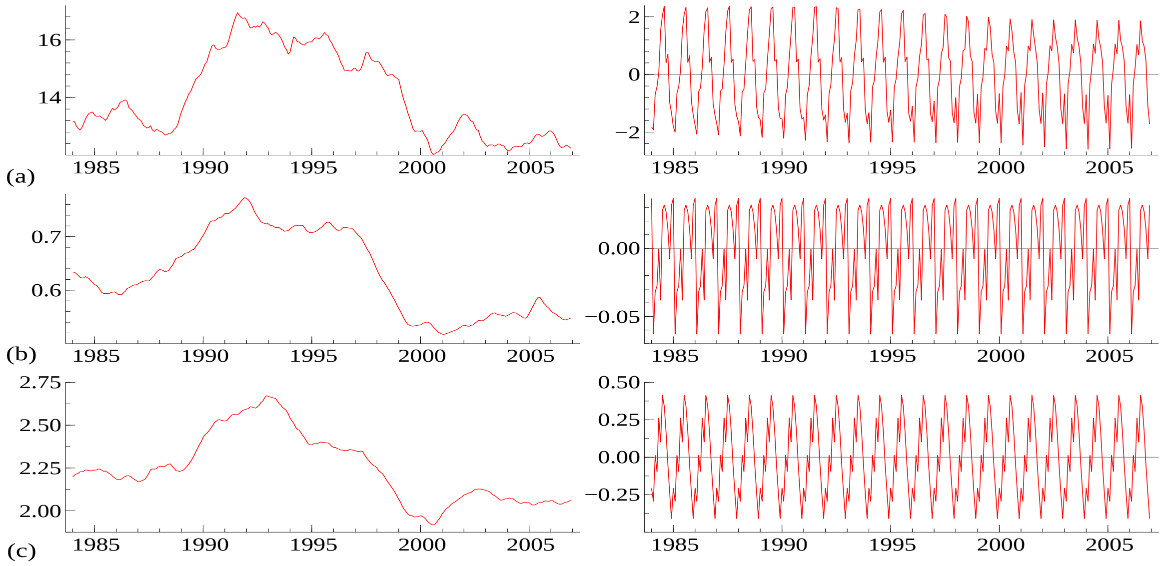

Figure 2.12: Trend and seasonal components in violent crime rates: (a) aggravated assault; (b) murder; (c) rape.

and Koopman (2001). Here we consider the auxiliary residuals for the level component to detect possible level breaks in the time series. It may enable us to find unusual events in the patterns of the analysed crime series, that would indicate the time position for the intervention analysis. We fitted the univariate STS models presented in Section 2.5 and extracted the auxiliary residuals. In Figures 2.13 and 2.14 we plotted the number of positive values of auxiliary residuals associated with the levels of fitted local level and local linear trend models. These positive values are counted for each year in the sample and plotted against time. Visual analysis of these positive counts suggests the following:

- Burglary: consecutive small positive values in 1998-2000 suggest a downward trend.
- Larceny: two consecutive years at the end of the 1990s with small positive values.
- Motor vehicle theft: we do not have a very clear picture.
- Robbery: consecutive small positive values in 1997-1999 suggest a downward trend.
- Aggravated assault: a more volatile series but there is only in the years 1998-99 consecutive numbers with low counts.
- Murder: it seems that from 1992 to 2000 there is a hole in counts of positive residuals.
- Rape: it seems that from 1993 to 2000 there is a hole in counts of positive residuals.

The low counts in the violent crime series (aggravated assault, murder, and rape) in the nineties are "significant" (less than 3), according to the non-parametric sign test. In the aftermath of the policy changes in the mid-1990s, we have found some evidence that in the final years of mid-1990s the trend shocks are negative in a consecutive number of years, leading to a strong change in the trend. Of course, there may be other reasons for these negative shocks in these years but it may be reasonable also to attribute the law/policy change in the mid-1990s as a reason for the negative shocks in the later years.

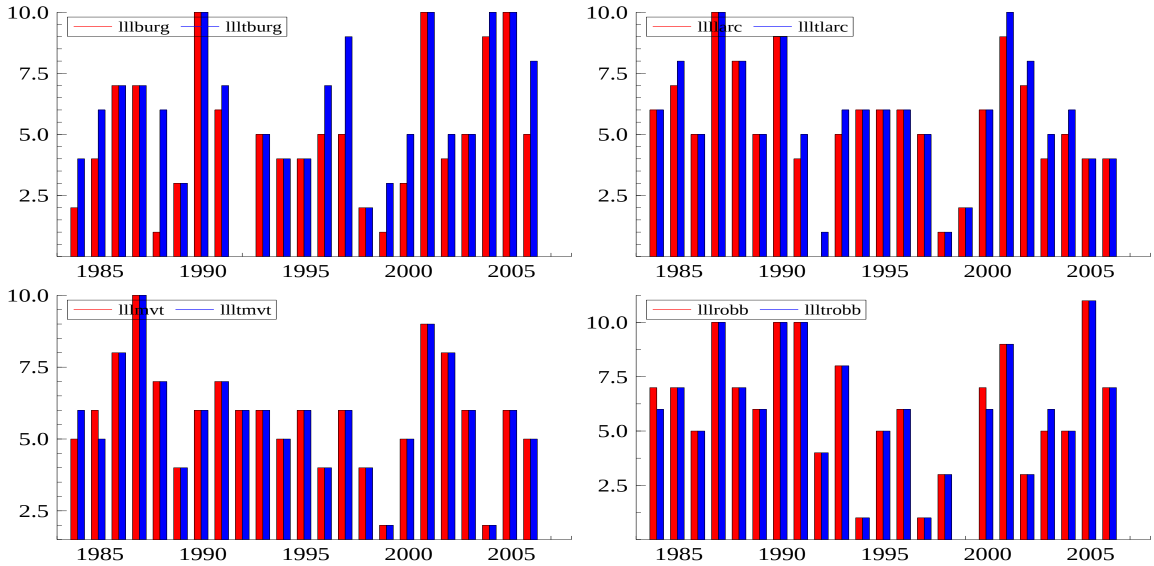

Figure 2.13: Auxiliary residual plots of the LL and LLT fitted models: Property crimes

STS intervention analysis results

After diagnostic checking, we fitted a local linear trend model as defined in (2.16), for burglary and larceny series, controlling for the outlier in December 1989. For other analysed crime rate series we fitted a local level model as defined in (2.15). All series are modelled with a seasonal component. Burglary, larceny, robbery, and aggravated assault are modelled with a stochastic seasonal, whereas motor vehicle theft, murder and rape are modelled with a deterministic seasonal. The estimation results for the structural time series model, separately for property and violent crimes, are presented in Tables 2.8 and 2.9. Negative estimation effects of the new legislation are found for burglary, larceny, murder, and rape. However, statistically significant negative impact of the new legislation is only found for burglary and murder. Positive estimation effects of the new legislation are found for motor vehicle theft, robbery, and aggravated assault, but the estimation results are not statistically significant. In comparison to other estimation results presented, both regression models with ARMA errors and the STS models indicate that the new legislation negatively affected burglary and murder time series, but not so much other analysed crime rates. The magnitude of the estimated coefficients is somewhat smaller for property crimes in the STS models and we found positive effect for the robbery series, although not statistically significant.

Considering our analysis, we conclude that the new legislation only affected burglary and murder offences, but not so much other crime categories. In Section 2.6, we are

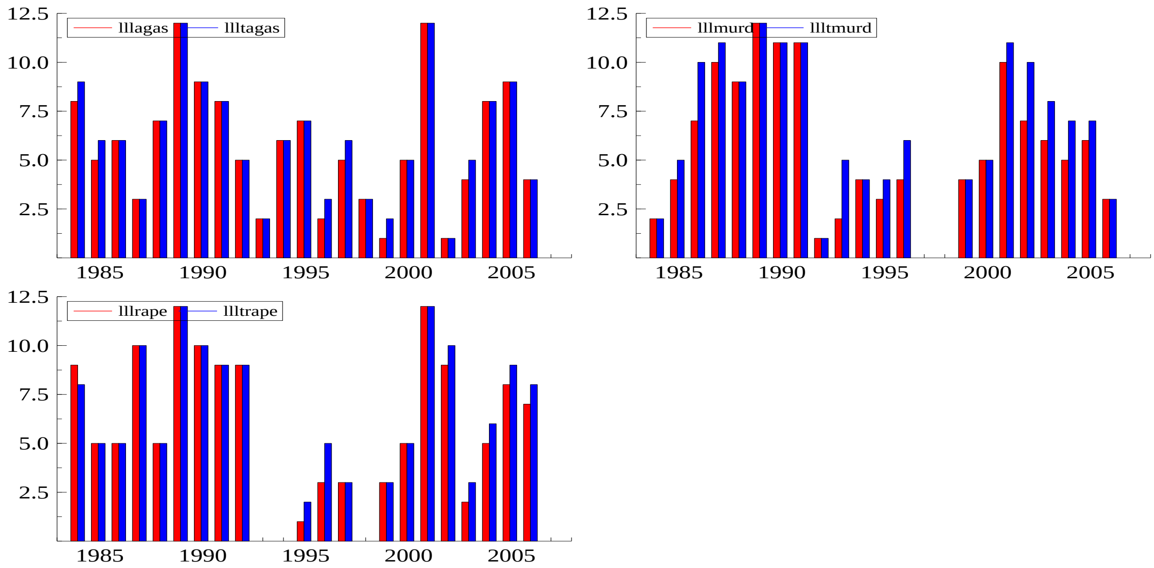

Figure 2.14: Auxiliary residual plots of the LL and LLT fitted models: Violent crimes

going to consider burglary and murder as 'eligible' crimes and others as 'not eligible' crimes. The new legislation also targeted robbery and rape offences, hence we consider these crimes as 'eligible' as well. Larceny, motor vehicle theft, and aggravated assaults are considered as 'non eligible' crimes. Following the methodology of Harvey (1996), we are going to estimate bivariate structural time series models with control groups, where burglary, robbery, murder, and rape crimes are considered as "treatment" series, while larceny, motor vehicle theft, and aggravated assault are considered as "control" series.

STS smooth step intervention analysis results

The results from the STS approach with intervention variable B_t defined in (2.2) confirm the preliminary findings of a deterrent impact of the new legislation for crimes, in particular those with a violent nature. All evidence is based on the intervention variable B_t for which the impact starts in January 1995. In order to investigate the robustness of these results and whether the impact was abrupt or more gradual, in this section we repeat the empirical analyses of the previous section on the basis of S_t instead of B_t. The gradual break S_t is defined in (2.3), it starts in January 1995 and we let the gradual break end in different years. In Table 2.10 the t-statistic (absolute value) of the estimated interventions based on B_t and S_t (for different lengths of the gradual break) are reported.

We find some clear evidence that a gradual break (S_t) has been more likely than an abrupt break (B_t) in 1995. In particular, the gradual break that ends in 2000 shows a significant effect for many crime series. Almost all violent crime series are significantly affected by the gradual break (1995–2000). Also the more violent property crimes like larceny and robbery are significantly affected by the gradual break. The estimated coefficients associated with this gradual break S_t for all crime series are found to be negative.

We should however mention that more gradual breaks also lead to more general in-

Table 2.8: Estimated interventions for structural time series models - Property crimes

	Statistic	Burglary			Larceny			Motor vehicle theft		Robbery	
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(1)	(2)
		LLT	LLT	LLT	LLT	LLT	LLT	LL	LL	LL	LL
Intervention				-3.96			-8.04		1.59		0.48
(95:01)				(-1.82)			(-1.17)		(1.22)		(0.78)
Outlier			-12.63	-12.56		-37.05	-36.93				
(89:12)			(-5.55)	(-5.54)		(-5.70)	(-5.69)				
Variances											
	σ^2_{irr}	3.30	2.59	2.57	25.83	19.90	19.80	0.87	0.96	0.30	0.29
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
	σ^2_{lvl}	1.38	1.34	1.35	15.86	17.05	17.32	0.66	0.62	0.09	0.09
		[0.42]	[0.52]	[0.52]	[0.61]	[0.86]	[0.87]	[0.75]	[0.68]	[0.31]	[0.31]
	σ^2_{seas}	0.01	0.01	0.01	0.05	0.05	0.05			0.00	0.00
		[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]			[0.00]	[0.00]
Diagnostics											
Independence	$Q(15)$	14.93	19.98	19.62	21.33	24.42	23.01	29.81	28.69	19.32	19.13
First-order ACF	$r(1)$	0.08	0.10	0.09	0.01	0.03	0.03	0.08	0.08	0.10	0.09
Homoscedasticity	$H(.)$	0.43	0.57	0.57	0.83	1.13	1.12	0.62	0.63	0.65	0.65
Normality	N	29.82	2.77	1.08	47.91	7.94	7.10	1.35	1.25	4.61	4.39
Goodness-of-fit											
	LogL	-307.07	-292.66	-290.73	-586.79	-570.45	-568.34	-114.67	-114.19	29.01	28.33
	p.e.v.	8.52	7.69	7.59	71.06	63.55	63.18	1.95	1.94	0.68	0.68
	R^2_s	0.16	0.24	0.25	0.20	0.29	0.29	0.11	0.11	0.19	0.19
	AIC	2.25	2.16	2.15	4.37	4.27	4.27	0.76	0.77	-0.28	-0.28

Note: Sample size = 276; t-statistic in round brackets; q-ratio (ratio of the estimated standard deviations of the state disturbances and estimated standard deviation of the irregular) in square brackets; MVT fitted with a deterministic seasonal.

Table 2.9: Estimated interventions for structural time series models - Violent crimes

	Statistic	Aggravated assault		Murder		Rape	
		(1)	(2)	(1)	(2)	(1)	(2)
		LL	LL	LL	LL	LL	LL
Intervention			0.13		-0.10		-0.15
(95:01)			(0.20)		(-1.96)		(-1.24)
Variances							
	σ^2_{irr}	0.51	0.51	0.00	0.01	0.05	0.05
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
	σ^2_{lvl}	0.06	0.07	0.00	0.00	0.00	0.00
		[0.13]	[0.13]	[0.02]	[0.02]	[0.02]	[0.03]
	σ^2_{seas}	0.00	0.00				
		[0.00]	[0.00]				
Diagnostics							
Independence	$Q(15)$	10.46	10.55	18.80	17.65	8.15	7.25
First-order ACF	$r(1)$	0.07	0.07	0.05	0.05	0.01	0.02
Homoscedasticity	$H(.)$	0.72	0.72	0.60	0.60	0.63	0.62
Normality	N	6.02	6.13	5.46	5.64	6.23	5.92
Goodness-of-fit							
	LogL	-3.36	-4.30	574.52	572.91	363.76	361.87
	p.e.v.	0.86	1.03	0.01	0.01	0.05	0.05
	R^2_s	0.30	0.30	0.40	0.40	0.41	0.42
	AIC	-0.05	-0.04	-4.45	-4.46	-2.86	-2.85

Note: Sample size = 276; t-statistic in round brackets; q-ratio (ratio of the estimated standard deviations of the state disturbances and estimated standard deviation of the irregular) in square brackets. Murder and rape fitted with a deterministic seasonal.

terpretations of the break. A more smooth break affects the overall trend more generally in the period 1995–2000 and therefore cannot be exclusively associated with an event in, say, January 1995. However, in our empirical study, the longer prison sentences imposed under the 1994 sentencing reform are likely to have gradual effect on Virginia's prison population. For example, when we consider the inmates confined for offences committed in early 1996, about 25% of this new inmate popolation admitted to the prison came in under the old parole system, before 1995, while at the end of 2000 this number was about 1%. This provides some justification that the gradual intervention S_t should be considered in our study.

2.6 The multivariate structural time series approach

2.6.1 The multivariate STS approach: state space formulation

Until now, we have discussed univariate structural time series models, which means that we modelled only one time series at the time. In the case of structural time series models, we can easily generalise the analysis of one time series to the simultaneous analysis of two or more time series (say p). The multivariate time series model can also be represented by the state space form

$$y_t = Z_t\alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, H_t) \tag{2.39}$$

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t, \qquad \eta_t \sim \mathcal{NID}(0, Q_t) \tag{2.40}$$

for $t = 1, \ldots, n$. The observation or measurement equation (2.39) specifies $p \times 1$ vector y_t containing the values of p observed time series at time point t. The $p \times 1$ irregular vector ε_t contains the p observation disturbances, one for each time series in y_t. The p observation disturbances are assumed to have zero means and an unknown variance-covariance structure represented by a variance matrix H_t of order $p \times p$. The $m \times 1$ state vector α_t contains unobserved variables and unknown fixed effects. Matrix Z_t of order $p \times m$ links the unobservable factors and regression effects of the state vector with the observation vector. Matrix T_t in (2.40) is called the transition matrix of order $m \times m$. The $r \times 1$ vector η_t contains the state disturbances with zero means and unknown variances and covariances collected in the variance matrix Q_t of order $r \times r$. In many standard cases, $r = m$ and matrix R_t is the identity matrix I_m. In other cases, matrix R_t is an $m \times r$ selection matrix with $r < m$. Although matrix R_t can be specified freely, it is often composed of a selection from the r columns of the identity matrix I_m.

In the multivariate framework the trend, seasonal, and irregular components have become vectors. Further, the disturbances associated with the components are vectors with variance matrices. These extensions imply that trends and seasonals of individual series can be correlated. For example, the trend of one series also applies to another series after appropriate scaling. When correlations are high, it means that components will be estimated with the combined use of more time series. Hence, a more precise estimate of the unobserved trend is obtained as a result. In the limiting case of perfect correlations (equal to one) between trends of individual series, the trend component is an equally weighted sum of the individual series. In the perfect correlation case, the trend is said to

Table 2.10: Estimated interventions from structural time series models for both property and violent crimes

Offence	break 1995	gradual 1996	gradual 1998	gradual 2000	gradual 2002	gradual 2004	gradual 2006
	B_t	S_t	S_t	S_t	S_t	S_t	S_t
Burglary	1.82	0.36	0.17	1.59	0.62	1.28	0.04
Larceny	1.17	0.24	0.22	2.37	1.25	1.36	1.59
Motor vehicle theft	1.22	0.07	0.21	1.19	0.96	0.99	1.71
Robbery	0.78	0.06	0.57	2.32	1.18	1.46	0.73
Aggravated assault	0.20	0.09	0.31	1.75	1.13	1.77	1.62
Murder (2nd degree)	1.96	0.40	1.15	3.58	2.17	1.23	0.94
Rape (forcible)	1.24	0.53	0.81	2.36	1.11	0.99	0.81

Note: We report t-tests (absolute values) for the step B_t and smooth step S_t interventions, see (2.2) and (2.3) respectively. The break for B_t takes place in January 1995. The start of the smooth break S_t is τ_1 and corresponds with January 1995 while τ_2 is the end of the gradual break and is January in the year indicated by the column headings.

be common. The same argument applies for the seasonal component. Explanatory and intervention variables can be added in the same way as for a univariate regression models.

The general framework of a state space model can be used for multivariate time series analysis as well. As an example, we consider a case with $p = 2$ and with vectors and matrices given by

$$\alpha_t = \begin{pmatrix} \mu_t^{(1)} \\ \nu_t^{(1)} \\ \beta_t^{(1)} \\ \mu_t^{(2)} \\ \nu_t^{(2)} \\ \beta_t^{(2)} \end{pmatrix}, \qquad \eta_t = \begin{pmatrix} \xi_t^{(1)} \\ \zeta_t^{(1)} \\ \xi_t^{(2)} \\ \zeta_t^{(2)} \end{pmatrix}, \qquad T_t = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$R_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$Z_t = \begin{bmatrix} 1 & 0 & x_t & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & x_t \end{bmatrix}, \qquad H_t = \begin{bmatrix} \sigma^2_{\varepsilon^{(1)}} & \mathrm{cov}(\varepsilon^{(1)}, \varepsilon^{(2)}) \\ \mathrm{cov}(\varepsilon^{(1)}, \varepsilon^{(2)}) & \sigma^2_{\varepsilon^{(2)}} \end{bmatrix}, \text{ and}$$

$$Q_t = \begin{bmatrix} \sigma^2_{\xi^{(1)}} & 0 & \mathrm{cov}(\xi^{(1)}, \xi^{(2)}) & 0 \\ 0 & \sigma^2_{\zeta^{(1)}} & 0 & \mathrm{cov}(\zeta^{(1)}, \zeta^{(2)}) \\ \mathrm{cov}(\xi^{(1)}, \xi^{(2)}) & 0 & \sigma^2_{\xi^{(2)}} & 0 \\ 0 & \mathrm{cov}(\zeta^{(1)}, \zeta^{(2)}) & 0 & \sigma^2_{\zeta^{(2)}} \end{bmatrix}.$$

These matrices imply that a multivariate trend model is a bivariate local linear model, with the same explanatory variable x_t applied to both series in y_t. The superscripts (1) and (2) in the matrices and vectors imply whether a particular component or a variance belong to the first or to the second series, respectively. In the equation format, we get two observation equations corresponding to the matrix notation in (2.39) and six state equations corresponding to the matrix notation in (2.40):

$$\begin{aligned}
y_t^{(1)} &= \mu_t^{(1)} + \beta_t^{(1)} x_t + \varepsilon_t^{(1)}, \\
y_t^{(2)} &= \mu_t^{(2)} + \beta_t^{(2)} x_t + \varepsilon_t^{(2)}, \\
\mu_{t+1}^{(1)} &= \mu_t^{(1)} + \nu_t^{(1)} + \xi_t^{(1)}, \\
\nu_{t+1}^{(1)} &= \nu_t^{(1)} + \zeta_t^{(1)}, \\
\beta_{t+1}^{(1)} &= \beta_t^{(1)}, \\
\mu_{t+1}^{(2)} &= \mu_t^{(2)} + \nu_t^{(2)} + \xi_t^{(2)}, \\
\nu_{t+1}^{(2)} &= \nu_t^{(2)} + \zeta_t^{(2)}, \\
\beta_{t+1}^{(2)} &= \beta_t^{(2)}.
\end{aligned}$$

In the example above, the same model is applied to the two time series in y_t. However, it is also possible to use two different state space models for the two series. As we shall demonstrate in our estimation results, it is possible to include the intervention variable only in the first equation $y_t^{(1)}$ and not in the second equation $y_t^{(2)}$. In some applications, it is also convenient to have matrix Q_t as a block diagonal matrix, without any alternations to the underlying model. For a matrix representation of the last two cases, we refer the interested reader to Commandeur and Koopman (2007).

In a multivariate state space analysis, the disturbances of the observation and state equations can be correlated. For example, if the disturbances in the slope components are uncorrelated, that is $\mathrm{cov}(\zeta^{(1)}, \zeta^{(2)}) = 0$, the slope components are independent. On the other hand, if the disturbances in the slope components are correlated, that is $\mathrm{cov}(\zeta^{(1)}, \zeta^{(2)}) \neq 0$, the multivariate time series model is referred to as a *seemingly unrelated time series equation model.* The level of dependence is measured by the correlation between the two disturbances as given by

$$\mathrm{corr}(\zeta^{(1)}, \zeta^{(2)}) = \frac{\mathrm{cov}(\zeta^{(1)}, \zeta^{(2)})}{\sqrt{\sigma^2_{\zeta^{(1)}}\sigma^2_{\zeta^{(2)}}}},$$

where $-1 \leq \mathrm{corr}(\zeta^{(1)}, \zeta^{(2)}) \leq 1$. In the extreme case of $\mathrm{corr}(\zeta^{(1)}, \zeta^{(2)}) = \pm 1$, a particular slope component can be expressed as a linear combination of the other slope component. For example, if the slope disturbances are perfectly correlated, we have $\nu_t^{(2)} = a + b\nu_t^{(1)}$. In this case, slope components are said to be *common.* In the case of $-1 < \mathrm{corr}(\zeta^{(1)}, \zeta^{(2)}) < 1$, the variance matrix

$$\begin{bmatrix} \sigma^2_{\zeta^{(1)}} & \mathrm{cov}(\zeta^{(1)}, \zeta^{(2)}) \\ \mathrm{cov}(\zeta^{(1)}, \zeta^{(2)}) & \sigma^2_{\zeta^{(2)}} \end{bmatrix},$$

has rank 2. In the case of $\mathrm{corr}(\zeta^{(1)}, \zeta^{(2)}) = \pm 1$, the rank of this variance matrix equals 1. Therefore, the rank of the variance matrix determines whether components are common. For multivariate models with $p > 2$ and a variance matrix with $q > 0$, the number of common components is equal to q and the number of rank restrictions is $r = p - q$. This type of analysis is known as *factor analysis* or *principal components analysis.* When $r = p - q$ rank restrictions are exercised, the p slope components are the result of linear combinations of q common slope components. The same argument can be applied to the disturbances of other components.

In general, a variance matrix is unknown and needs to be estimated. The estimated coefficients determine the rank of the matrix and the nature of the relationship between the individual elements of the component vector. In some cases, it might be necessary to impose rank restrictions. The rank of a particular variance matrix can be imposed by considering the decomposition of a symmetric positive semi-definite matrix such as

$$\begin{bmatrix} \sigma^2_{\zeta^{(1)}} & \mathrm{cov}(\zeta^{(1)}, \zeta^{(2)}) \\ \mathrm{cov}(\zeta^{(1)}, \zeta^{(2)}) & \sigma^2_{\zeta^{(2)}} \end{bmatrix} = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{bmatrix} a & b \\ 0 & c \end{bmatrix},$$

with coefficients $a, c \geq 0$.

The issue of common level and slopes is important, since it is often interesting to find the common behaviour between the different time series in a multivariate time series analysis. For example, finding common factors yields models with more efficient inferences and forecasts. An important feature of this model is that the common nonstationary levels can be removed by a linear combination of the two series. This property is known in the econometrics literature as *cointegration*.

Finally, if the variance matrices H_t and Q_t in (2.39) and (2.40) are *diagonal*, the rows of Z_t are orthogonal, and T_t is appropriately chosen, we actually carry out p separate univariate analyses. In this case, we have a 'really unrelated' time series equations model.

2.6.2 Multivariate STS intervention analysis results

The multivariate structural time series model for crime series can be used to assess the effect of parole abolition and reformed sentencing in Virginia. The results can be more convincing than the results from an univariate state space approach because more time series are involved simultaneously in the analysis. Following the methodology of (Harvey 1996), we estimate bivariate structural time series models with control groups. We divide the crime categories into two groups: the more serious crimes targeted by the new legislation are considered as a "treatment" group or crimes eligible for enhancements under the new legislation (burglary, robbery, murder, and rape) and the less serious felonies, that are not affected much by the new legislation are considered as a "control" group or crimes that are not eligible for enhancements under the new legislation (larceny, motor vehicle theft, and aggravated assault). Considering our previous estimation results, where burglary has always been found significantly affected by the new legislation, and considering that the new law targeted murder, rape, and robbery in particular (see Figure 2.1), this seems like a plausible assumption. Therefore, if we can show that the treatment series were significantly affected by the new legislation, while the control series were not affected by the intervention, we have an even stronger case in favour or against the effect of this legislation.

What is an economic interpretation for having "treatment" and "control" crime groups? Observed changes in crime around the time of the introduction of the new legislation "may reflect a combination of the true deterrent impact of harsher repeat-offender enhancements and of other factors correlated with but not caused by the law change, such as changes in demographics, in other state policies, and in broad social norms against crime," Kessler and Levitt (1999).

The estimation results for burglary as a "treatment" series and three different "control" series are presented in Table 2.11. The similar specifications for robbery, murder, and rape as "treatment" series are presented in Tables 2.11, 2.13, and 2.14, consecutively. We find significant negative effects in two out three bivariate STS specifications for burglary, with motor vehicle theft and aggravated assault as control groups. The estimated coefficient ranges from -4 to -5, which is in the neighbourhood of a univariate estimate of -4. The estimated bivariate models satisfy well the diagnostic requirements, although none give a superior specification for burglary when compared to the preferred univariate model. Looking at the Akaike information criterion, in none of the multivariate regressions for burglary does it go down to 2.15 as in the preferred univariate specification.

A similar picture is obtained for the bivariate STS specifications for murder series (Table 2.13). Negative significant effect of the new legislation are found in all three bivariate STS models with control groups. When motor vehicle theft and aggravated assault are treated as non-eligible crimes, the estimated effect of the new legislation on murder is -0.10, which is the same as the univariate result. The estimated bivariate models satisfy well the diagnostic requirements, except for the model with larceny as a control group (estimated effect is -0.16), because it seems that we do not control well for the serial correlation in the error term for murder ($Q = 322$). However, similar to bivarite models for burglary, none of the bivariate models fro murder give a superior specification when compared to the univariate model for murder. Looking at the Akaike information criterion, in none of the multivariate regressions for murder does it go down to -4.46 as in the univariate specification.

When we treat robbery as a treatment series, we do not find significant results in any of the three bivariate specifications. This finding suggests that although the new legislation targeted robbery as a most violent property crime, we do not find any confirmation that the behaviour of this series has been significantly altered by the new legislation. When we look at the bivarite results with rape as treatment series, we find a significant negative effect when rape is modelled together with aggravated assault as a control group. The estimated effect is -0.24, which is larger in absolute value from the univariate estimate of -0.15. However, the estimated AIC criteria (-2.71) does not suggest that the bivarite model is superior to the univariate model (AIC $= -2.86$).

Summarizing, the multivariate estimation results confirm that the new legislation significantly affected burglary (-4 to -5) and murder series (-0.10). When it comes to rape, we find a negative significant result (-0.24) only in one of our bivariate model specifications, which is not really superior to the univariate estimation result (-0.15). The simultaneous treatment of all seven crime series in one model does not lead to an improvement of our univariate or bivariate specifications. Also, we do not find any justification for a common trend between analysed crime series. In the next section we present a difference-in-difference estimation results, where we treat Virginia crime rates as eligible and non-eligible, and compare them with appropriate eligible and non-eligible crime rates in the rest of the United States.

2.7 General discussion of empirical findings

In this section we will present the difference-in-difference approach to the analysed crime series in the spirit of the paper by Kessler and Levitt (1999). It enables us to discuss our results in a wider perspective and to formulate some conclusions.

2.7.1 Relation to difference-in-difference approach

Since the new legislation in Virginia affected punishment levels for some crimes but not for others, we are able to evaluate its deterrent effects differing between "treatment" and "control" crime groups. In Sections 2.5 and 2.6, we have obtained our empirical results based on the time series methodology of (Harvey 1996). In order to check robustness

Table 2.11: Estimated interventions for multivariate STS models - Burglary as a treatment series

	Statistic	Burglary	Larceny	Burglary	MVT	Burglary	AA
Intervention		-2.72		-5.46		-3.96	
(95:01)		(-1.45)		(-2.72)		(-1.88)	
Outlier		-12.47	-38.29	-12.18		-11.90	
(89:12)		(-5.49)	(-5.90)	(-5.50)		(-5.49)	
Variances of disturbances							
	σ^2_{irr}	2.63	20.29	2.56	0.80	2.55	0.51
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
	σ^2_{lvl}	1.31	16.20	1.42	0.58	1.37	0.07
		[0.50]	[0.80]	[0.55]	[0.72]	[0.54]	[0.13]
	σ^2_{seas}	0.01	0.05	0.01	0.00	0.01	0.00
		[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Diagnostics							
Independence	$Q(15)$	19.54	22.35	18.94	24.09	20.63	10.54
First-order ACF	$r(1)$	0.10	0.02	0.09	0.05	0.10	0.07
Homoscedasticity	$H(.)$	0.55	1.11	0.54	0.53	0.59	0.70
Normality	N	1.10	11.41	1.69	2.94	1.05	5.08
Goodness-of-fit							
	LogL	-822.62		-398.11		-288.11	
	p.e.v.	7.57		7.49		7.59	
		10.93	62.59	1.14	1.89	0.69	0.86
	R^2_s	0.25	0.30	0.26	0.14	0.25	0.30
	AIC	2.29	4.40	2.27	0.90	2.29	0.11

Note: Sample size is 276; t-statistic in round brackets; q-ratio (ratio of the estimated standard deviations of the state disturbances and estimated standard deviation of the irregular) in square brackets; LLT model fitted to all specifications; MVT = Motor vehicle theft; AA = Aggravated assault.

Table 2.12: Estimated interventions for multivariate STS models - Robbery as a treatment series

	Statistic	Robbery	Larceny	Robbery	MVT	Robbery	AA
Intervention		0.64		-0.14		0.40	
(95:01)		(0.98)		(-0.25)		(0.68)	
Outlier			-33.45				
(89:12)			(5.34)				
Variances of disturbances							
	σ^2_{irr}	0.62	4.50	0.28	0.79	0.29	0.51
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
	σ^2_{lvl}	0.42	4.49	0.11	0.63	0.09	0.07
		[0.68]	[1.00]	[0.41]	[0.80]	[0.32]	[0.13]
	σ^2_{seas}	0.00	0.19	0.00	0.00	0.00	0.00
		[0.00]	[0.04]	[0.00]	[0.00]	[0.00]	[0.00]
Diagnostics							
Independence	$Q(15)$	44.89	24.69	14.47	25.64	18.18	11.99
First-order ACF	$r(1)$	0.05	0.03	0.05	0.08	0.09	0.08
Homoscedasticity	$H(.)$	0.96	1.17	0.59	0.60	0.65	0.70
Normality	N	3.63	10.49	6.76	3.26	5.26	5.61
Goodness-of-fit							
	LogL	-548.57		-81.67		24.57	
	p.e.v.	0.70		0.67		0.68	
		2.29	63.95	0.41	2.05	0.19	0.86
	R^2_s	0.16	0.28	0.20	0.07	0.19	0.30
	AIC	-0.09	4.42	-0.14	0.97	-0.13	0.10

Note: Sample size is 276; t-statistic in round brackets; q-ratio (ratio of the estimated standard deviations of the state disturbances and estimated standard deviation of the irregular) in square brackets; LLT model fitted to all specifications; MVT = Motor vehicle theft; AA = Aggravated assault.

Table 2.13: Estimated interventions for multivariate STS models - Murder as a treatment series

	Statistic	Murder	Larceny	Murder	MVT	Murder	AA
Intervention		-0.16		-0.10		-0.10	
(95:01)		(-5.65)		(-2.16)		(-2.25)	
Outlier			-37.04				
(89:12)			(-5.71)				
Variances of disturbances							
	σ^2_{irr}	0.01	19.81	0.01	0.79	0.01	0.51
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
	σ^2_{lvl}	0.00	17.13	0.00	0.63	0.00	0.07
		[0.00]	[0.86]	[0.02]	[0.79]	[0.02]	[0.13]
	σ^2_{seas}	0.00	0.05	0.00	0.00	0.00	0.00
		[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Diagnostics							
Independence	$Q(15)$	256.68	24.43	12.67	27.00	15.04	10.48
First-order ACF	$r(1)$	0.32	0.03	0.03	0.08	0.06	0.06
Homoscedasticity	$H(.)$	1.10	1.13	0.58	0.59	0.61	0.71
Normality	N	9.40	8.04	5.04	3.46	3.17	5.54
Goodness-of-fit							
	LogL	-55.06		435.07		552.04	
	p.e.v.	0.01		0.01		0.01	
		-0.01	63.55	0.01	2.06	0.01	0.86
	R^2_s	0.20	0.29	0.41	0.06	0.41	0.30
	AIC	-4.01	4.41	-4.32	0.98	-4.32	0.10

Note: Sample size is 276; t-statistic in round brackets; q-ratio (ratio of the estimated standard deviations of the state disturbances and estimated standard deviation of the irregular) in square brackets; LLT model fitted to all specifications; MVT = Motor vehicle theft; AA = Aggravated assault.

Table 2.14: Estimated interventions for multivariate STS models - Rape as a treatment series

	Statistic	Rape	Larceny	Rape	MVT	Rape	AA
Intervention		-0.19		-0.20		-0.24	
(95:01)		(-1.79)		(-1.67)		(-2.52)	
Outlier			-35.28				
(89:12)			(-5.52)				
Variances of disturbances							
	σ^2_{irr}	0.05	20.05	0.05	0.80	0.05	0.52
		[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
	σ^2_{lvl}	0.00	16.89	0.00	0.62	0.00	0.06
		[0.02]	[0.84]	[0.03]	[0.77]	[0.02]	[0.12]
	σ^2_{seas}	0.00	0.05	0.00	0.00	0.00	0.00
		[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Diagnostics							
Independence	$Q(15)$	6.95	23.62	7.89	26.27	6.11	10.16
First-order ACF	$r(1)$	0.02	0.04	-0.00	0.07	0.02	0.07
Homoscedasticity	$H(.)$	0.62	1.13	0.64	0.58	0.67	0.70
Normality	N	2.43	10.76	2.25	3.07	1.84	4.77
Goodness-of-fit							
	LogL	-216.52		226.01		343.59	
	p.e.v.	0.05		0.05		0.05	
		0.34	63.38	0.05	2.06	0.03	0.85
	R^2_s	0.42	0.29	0.42	0.06	0.42	0.30
	AIC	-2.71	4.41	-2.71	0.98	-2.71	0.09

Note: Sample size is 276; t-statistic in round brackets; q-ratio (ratio of the estimated standard deviations of the state disturbances and estimated standard deviation of the irregular) in square brackets; LLT model fitted to all specifications; MVT = Motor vehicle theft; AA = Aggravated assault.

of our estimation results, we also follow the method of Kessler and Levitt (1999), who used a difference-in-difference approach to evaluate the deterrent effect of repeat-offender enhancements in California in the United States. Table 2.15 shows data for years 1991 to 2003, for Virginia and the rest of the United States. Hence, we look at 4 years of data before the change in the legislation and 8 years of data after the new legislation. Inspection of Table 2.15 reveals that levels of crime in Virginia are lower than those in the rest of the nation, but increases and decreases in the crime rates in Virginia parallel those in the rest of the United States. In the years before the passage of the law, eligible crimes were already on its downward slope, and they seemed to have been falling further after the 1995.

Table 2.16 shows the results of the simple difference-in-difference exercise. Eligible crimes in Virginia are the "treatment" group, while ineligible crimes in Virginia are the "control" group. Additionally, eligible and ineligible crime categories in the rest of the United States are also presented. The eligible and ineligible crimes outside of Virginia have not been affected by the new legislation and thus provide another control group. The first two columns of the table 2.16 give percent changes in crime rates before passage of the law, while the final four columns show crime patterns after the law change. Both eligible and ineligible crimes seem to have been falling after the law change in 1995. Row 3 of Table 2.16 computes the difference between patterns in eligible and ineligible crimes for each time period in Virginia. As we have seen in our descriptive analysis in Section 2.3, in the years before the passage of the law, eligible crimes were already on its downward slope. After the new law, there is in immediate 3% reduction in eligible crimes that grows over time to 11% in 2001. This approach provides one estimate of the effect of the new law on the eligible crimes. The rest of the United States experience similar changes in crime rates after 1995, but with a smaller magnitude.

The most convincing estimate of the effect of the new legislation is the change in eligible crimes relative to ineligible crimes in Virginia minus the corresponding change outside of Virginia (a "difference-in-difference" estimator). These results are presented at the bottom of Table 2.16. Before the new legislation, crime rates in the rest of the United States were falling at the higher rate than those in Virginia. After 1995, an immediate decline of 0.1% in eligible crimes occurs. Four years after the law change, the decline has become much larger, at 7.2%. These results confirm that crimes in the eligible categories responded immediately to the passage of the new legislation. We see not only an immediate fall in crime with the adoption of the new legislation, but also a continued decline in crime rates over several years after the passage of the new legislation. We could say that the deterrence (immediate) effect of the new legislation was about 0.1% and over the longer run as incapacitation and further deterrence kicked in, the effect of the new legislation was around 7%.

2.7.2 Review of empirical findings

In this chapter we adopted different time series methodologies to investigate the impact of parole abolition and sentence reform in Virginia on reported crime rates. The Commonwealth of Virginia abolished parole and reformed sentencing for all felony offences committed on or after January 1, 1995. To examine the impact of Virginia's new legis-

Table 2.15: Crime rates in Virginia and the rest of the United States, 1991-2003

	1991	1993	1995	1997	1999	2001	2003
Virginia							
Eligible crimes							
Murder	9.3	8.3	8.6	7.2	5.7	5.1	5.6
Rape	29.9	32.1	27.2	27.0	25.0	24.6	24
Robbery	137.6	142.0	131.7	124.5	101.1	95.4	90.3
Burglary	783.1	667.7	595.2	571.4	471.6	439.7	391.5
Noneligible crimes							
Aggravated assault	196.4	189.8	194.9	186.5	182.8	166.2	155.9
Motor vehicle theft	338.6	285.5	292.6	280.5	261.2	262.1	242.5
Larceny	3112.5	2790.1	2740.0	2679.0	2326.5	2185.1	2070.0
US, excluding Virginia							
Eligible crimes							
Murder	9.8	9.5	8.2	6.8	5.7	5.6	5.7
Rape	42.6	41.4	37.3	36.1	33.0	32.0	32.3
Robbery	276.2	258.9	223.2	187.8	151.4	149.9	143.5
Burglary	1264.1	1110.8	997.1	927.8	778.1	748.6	749.6
Noneligible crimes							
Aggravated assault	439.4	447.0	424.0	387.1	338.3	322.5	298.6
Motor vehicle theft	667.2	614.6	567.2	511.5	426.7	435.0	438.4
Larceny	3232.1	3040.2	3051.1	2897.3	2556.5	2543.3	2423.5

Note: Values in the table are reported crime rates per 100,000 residents. Data for Virginia and the rest of the United States are from *Crime in the Nation*, the Uniform Crime Reports published annually by the Federal Bureau of Investigation.

Table 2.16: Estimates of the impact of the parole abolition and sentence reform on Virginia crime rates

	pre-1995		post-1995			
Geographic regrion & crime category	1991-1995	1993-1995	1995-1997	1995-1999	1995-2001	1995-2003
Virginia						
Eligible crime	-11.2	-7.4	-6.6	-21.5	-26.0	-28.1
Non-eligible crimes	-8.8	1.1	-3.6	-10.7	-15.1	-20.5
VA eligible - VA non-eligible	-2.4	-8.5	-3.0	-10.8	-10.9	-7.6
Rest US						
Eligible crime	-17.3	-11.9	-10.8	-24.0	-25.9	-26.1
Non-eligible crimes	-8.0	-4.2	-7.9	-20.4	-21.3	-24.3
Rest US eligible - Rest US non-eligible	-9.3	-7.7	-2.9	-3.6	-4.6	-1.8
Diff-in-diff						
(VA eligible - VA non-eligible) -						
(Rest US eligible - Rest US non-eligible)	6.9	-0.8	-0.1	-7.2	-6.3	-5.8

Note: Table entries are average percentage changes in crime rates per 100,000 residents over the relevant crime categories in the years listed. Crimes eligible for parole abolition and sentence reform in Virginia are murder, rape, robbery and burglary. Ineligible crimes are aggravated assault, motor vehicle theft, and larceny. Values in the third row are the difference between rows 1 and 2. Values in the sixth row are the difference between rows 4 and 5. Values in the bottom row are the difference between rows 3 and 6. The reform took effect in January 1995.

lation on reported crime rates, we considered different empirical approaches to the intervention analysis, such as regression, ARIMA, and structural time series approaches. The STS models are estimated both in the univariate and multivariate dimensions. In order to compare our findings with the existing literature, we have also applied a difference-in-difference estimation approach. Examined crime rate series include burglary, larceny, motor vehicle theft, robbery, aggravated assault, murder, and rape.

We recognize the following contributions of our chapter. First, we see it as a comprehensive evaluation of the new legislation. According to the Virginia crime officials, research to date was unable to determine if the observed reductions in crime rates were due to specific anti-crime initiatives. Hence, this chapter tries to contribute to better understanding of the statistical relationship between anti-crime efforts and crime reductions over time.

Second, we view our work as an introduction of STS methods to crime data analysis. To our knowledge, the STS estimation approach has not been used extensively in crime analysis. On the other hand, the regression and especially the ARIMA models have been widely used in the criminal justice literature. Suggestion for the application of the Kalman filter in the crime economics literature has been made by Britt (2001). Existing applications of the STS methodology in crime analysis are to be found in papers by Harvey and Fernandez (1989), Atkinson et al. (1997), Koopman et al. (2008), and Vujić and Koopman (2008). The structural time series methodology provides a broader framework and allows the parameters to change over the time period. This approach based on unobserved (possibly non-stationary) components does not require differencing which can results in messy time series when outlying observations are present. Moreover, the incorporation of explanatory and intervention variables is straightforward. Finally, the structural time series models are flexible and can be generalised to a multivariate setting without much extra effort. Hence, we view the present chapter as a potential contribution to time series methodology in criminology.

Third, when dealing with time series, we show superiority of the ARIMA and STS models in comparison to regression approach. Proposed models for analysing the intervention effects of parole abolition and sentence reform in Virginia clearly favour ARIMA and/or structural time series approaches to modelling intervention. Results using regression approaches are biased and the measured effects are not reliable because of the serially correlated errors. In addition to this, the intervention does not have to be obvious - there may be three sources of noise that might eliminate the intervention effect. These are trend, seasonality and random effects. Together with the fact that adjacent error terms tend to be correlated and that the proposed model has to account for this type of noise as well, ARIMA and STS models include all these effects and are more effective approaches in analysing time series intervention design. Once the sources of variance in the series have been controlled for, the impact of an intervention can be tested and measured with greater reliability. Therefore, we should concentrate on discussing the estimation results obtained using ARIMA and STS models.

The estimation results suggest that crimes in the eligible categories responded immediately to the passage of the new legislation. We see not only an immediate fall in crime with the adoption of the new legislation, but also a continued decline in crime rates over several years after the passage of the new legislation. We could say that the deterrence (immediate) effect of the new legislation was about -0.1% and over the longer run as incapacitation and further deterrence kicked in, the effect of the new legislation was around -7%. When it comes to the effect of the new legislation on the particular crime offences, we find that the new legislation significantly affected burglary (-4 to -5) and murder series (-0.10). In case of rape, we find a negative significant result (-0.24) only in one of our bivariate model specifications. This is not really superior to the univariate estimation result (-0.15) which is also not found to be significantly different from zero. The possible explanation for the non-significant effect on the aggravated assault would be the very method in which this category is reported. In contrast to other crimes, which are relatively well defined, "aggravated" assault requires discretion on the part of the police taking the report to distinguish it from "simple" assault. This way of reporting leaves room for the exercise of discretion and there is a possibility that the nature of this distinction has been changing over time (Blumstein, 2000). Hence, this might be a reason for the non-significant intervention coefficient of the aggravated assault which is, by definition, violent offence.

Although the new legislation targeted more violent than property crime offences, we do not find significant effects of the law change in Virginia for all serious felonies. The reasons might be the following ones. First, our modelling approach is not powerful enough to capture the effect of the new legislation. Table 2.16 suggests that the effect of the new legislation has been more gradual, spreading over several years between 1995 and 2003. However, even when we re-estimate our univariate STS results by introducing a logit transformation of the intervention dummy variable, we still do not find conclusive evidence. Second, other factors, such as economic conditions, might have affected observed patterns of recorded crime rates. For example, violent and property crime response differently to economic conditions. Property crimes are usually countercyclical, while violent crimes are not strongly affected by the economy. This means that including economic conditions in our research approach would have affected our property crime estimation results, but not so much estimated results on violent crimes. Third, the effect of other policies enacted in Virginia during this period might have affected our results. Looking at Table 2.2, we speculate that only "Virginia firearms transaction programme" in 1989 and "One handgun per 30-day purchase limit" in 1993 might have had significant effects. The first policy has been controlled for by introducing a blip dummy in 1989, while a blip dummy in 1993 has not been significant in any of the modelled crime series. Fourth, the new legislation has not been enforced by the prosecutors and judges in Virginia. According to Kessler and Levitt (1999), "in most states, three-strikes statues have never been enforced."[22] Further, Kessler and Levitt (1999) report that previous research on the impact of sentence enhancements on reported crime rates in the United States has reported conflicting findings and often find no effect (see for example Loftin et al., 1983).

Virginia's abolition of parole and reform of the sentencing system provides a useful

[22]During the 1990s, almost all states in the US adopted one or more mandatory sentencing laws. They all had one thing in common: they all imposed mandatory, statutory increases in prison sentences on individuals who were already going to be imprisoned. In addition, 24 states among which Virginia, enacted a new, more stringent breed of repeat-offender enhancements called "Three Strikes and You're Out" laws between 1993 and 1995, see for example Clark et al. (1997). In many states, three-strikes laws impose life imprisonment without parole for those convicted of three separate violent felonies.

social experiment to study. First the legislation was very sweeping and impacted all felonies. Further, such sweeping legislation was enacted at a time in which there were very large (and favourable) changes in a number of social and economic indicators. Finally, the 1990s also saw the implementation of a number of initiatives focused on reducing crime at the Federal, State and Community levels. Disentangling the impact of parole abolition from the other factors poses multiple design and analytical challenges.

2.A Appendix

2.A.1 Diagnostic and Goodness-of-Fit Tests

The significance tests of the presented estimation results are based on three assumptions concerning the residuals: independence, homoscedasticity, and normality. We list the tests used in the chapter in order to establish these properties of residuals. A very detailed discussion of the exact definition, computation, and interpretation of these diagnostic tests is given by Commandeur and Koopman (2007). We also list the goodness-of-fit measures used in order to choose the preferred model specification.

Diagnostic Tests

- The portmanteau Box-Ljung $Q(p)$ test statistic is used to check the assumption of independence of residuals. It is defined as a sum of the first p autocorrelation coefficients of the standardised regression residuals. If the residual autocorrelation for lag k is defined as

  $$r_k = \frac{\sum_{t=1}^{n-k}(e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^{n}(e_t - \bar{e})^2},$$

  where $\bar{e}$ is the mean of n residuals, then the Box-Ljung statistic is defined as

  $$Q(k) = n(n+2)\sum_{l=1}^{k}\frac{r_l^2}{n-l},$$

  for lags $l = 1, \ldots, k$. This should be tested against a χ^2 distribution with approximately $(k - w + 1)$ degrees of freedom, where w is the number of estimated hyperparameters (i.e., disturbance variances).

- The homoscedasticity of residuals is checked with H-statistic, which tests whether the variances of two successive residuals are equal to one another. It is computed as

  $$H(h) = \frac{\sum_{t=n-h+1}^{n} e_t^2}{\sum_{t=d+1}^{d+h} e_t^2},$$

  where d is the number of diffuse elements, and h is the nearest integer to $(n-d)/3$. This is typically a two-tailed tests and it should be tested against an F-distribution with (h, h) degrees of freedom.

- The normality of residuals is checked with the Bowman and Shenton normality test, N, which follows χ^2 distributed with two degrees of freedom. This is computed using the following statistic:

  $$N = n(\frac{S^2}{6} + \frac{(K-3)^2}{24}),$$

with

$$S = \frac{\frac{1}{n}\sum_{t=1}^{n}(e_t - \bar{e})^3}{\sqrt{(\frac{1}{n}\sum_{t=1}^{n}(e_t - \bar{e}))^3}}, \qquad K = \frac{\frac{1}{n}\sum_{t=1}^{n}(e_t - \bar{e})^4}{(\frac{1}{n}\sum_{t=1}^{n}(e_t - \bar{e}))^2}.$$

Null hypothesis here is that the data come from a Normal distribution. Sample from a Normal distribution have an expected skewness/kurtosis equal to zero. Any deviations from zero will increase the N statistic.

- $r(1)$ and $r(15)$ are the values of autocorrelations at lags 1 and 15.

Goodness-of-Fit Measures

- The one-step ahead prediction error variance (p.e.v.) is a one-step ahead prediction error variance. Tests used to establish whether the residuals of the state space models satisfy the assumptions of independence, homoscedasticity, and normality are applied to the standardised prediction errors, defined as:

$$e_t = \frac{\upsilon_t}{\sqrt{F_t}},$$

where υ_t are prediction errors, and F_t are their variances. The variance of the standardised prediction errors is approximately equal to one.

- The log-likelihood. The prediction errors υ_t and their variances F_t are also important for maximisation of the log-likelihood in state space methods. For univariate state space methods the diffuse log-likelihoods is defined as:

$$\text{Log}L_d = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=d+1}^{n}(\log F_t + \frac{\upsilon_t^2}{F_t}),$$

where d is the number of diffuse initial elements of the state. The value of the log-likelihood function is maximised by simultaneously minimising the prediction errors υ_t and their variances F_t. Similarly, in classical regression model parameter estimates are obtained by minimising the observation errors ε_t and their variances σ_ε^2.

- The R^2 value corrected for seasonal means (R_s^2). For seasonal data with a trend, the R^2 value corrected for seasonal means is a more appropriate measure of goodness-of-fit than the traditional coefficient of determination. This requires the sum of squares, SSDSM, obtained by substracting the seasonal means from the data's first differences (Δy_t). The coefficient of determination is then $R_s^2 = 1 - SSE/SSDSM$, where SSE is the residual sum of squares. Any model that has R_s^2 negative can be rejected (Harvey (1989)).

- The Akaike Information Criterion (AIC). In order to compare the different estimated models, we uses Akaike Information Criterion:

$$\text{AIC} = \frac{1}{n}[-2n\log L_d + 2(q+\omega)],$$

where n is the number of observations in the time series, $\log L_d$ is the value of the diffuse log-likelihood, q is the number of diffuse initial values in the state, and ω is the total number of disturbance variances estimated in the analysis. When we compared different models using the AIC criterion, models with smaller AIC are better fitted than models with larger AIC. AIC criterion compensates for the number of estimated parameters in a model and allows a fair comparison between models with different numbers of estimated parameters.

2.A.2 Definitions of the analysed crime categories

Source: Crime in Virginia, 2006

- *Aggravated Assault* - An unlawful attack by one person upon another wherein the offender uses a weapon or displays it in a threatening manner, for the purpose of inflicting severe or aggravated bodily injury. Simple assaults are excluded.

- *Burglary* - The unlawful entry into a building or other structure with the intent to commit a felony or a theft. Attempted forcible entry is included.

- *Larceny* - The unlawful taking of property from the possession of another person. Examples are pocket-picking, purse-snatching, shoplifting, theft from a building, theft of motor vehicle parts or accessories, etc. Attempted larcenies are included.

- *Motor Vehicle Theft* - The theft or attempted theft of a motor vehicle.

- *Murder* - The wilful (nonnegligent) killing of one human being by another.

- *Robbery* - The taking, or attempting to take, anything of value from the control, custody, or care of another person by force, or threat of force, or violence, and/or by putting the victim in fear of immediate harm.

- *Forcible Rape* - The carnal knowledge of a person, forcibly, and/or against that person's will; or not forcibly or against the person's will where the victim is incapable of giving consent because of his/her temporary or permanent mental or physical incapacity.

Chapter 3

Economic Trends and Cycles in Crime: A Study for England and Wales

This chapter is based on work done at the VU in 2003/2004, under the supervision of Prof. Siem Jan Koopman.

3.1 Introduction

By considering crime as a labour market phenomenon, a potential criminal weighs the expected gains against the expected costs and on the base of these decides whether or not to engage in crime. Ever since the seminal paper of Garry Becker (1968), crime economics relies on this behavioural model. Although the theoretical model is widely accepted among the social scientists, there are alternative strategies which attempt to empirically identify the causes of crime. These range from cross-sectional and panel data models, to macro level approaches that utilise aggregate time series data. Cross-section and panel data analysis of less geographically aggregated areas are an ideal tool for distinguishing between two alternative behavioural explanations for a link between crime and economic indicators (Levitt, 2001). On the other hand, time series data on a higher aggregation level, say national, are more effective in describing the long-run patterns and interactions with macroeconomic variables. Hence, the cross-section and panel data approaches might be a natural framework with which to analyse behavioural aspects of crime on an individual level, while the macro level approaches are more appropriate for examination of interaction between economic trends and crime.[1]

[1]Further point in favour of aggregate data approach is that at the level of individual police department or arrondissement, both administrative and political changes can lead to irregularities in reported data or to failures to report any data (Witt and Dryden Witte (1998) refer to papers by Donohue and Siegelman (1998) and Boggess and Bound (1993) in this respect). In such a setting, aggregate data can more convincingly reflect underlying trends than more disaggregated data. Additionally, in the process of aggregation, many individual peculiarities are averaged out and outliers are greatly reduced (Witt and Dryden Witte, 1998). Ralston (1999) cites the U.S. Congress which says that "the findings from research using annual aggregate data are fundamental to the appropriate legislation of national economic and social policy." In words of Dhiri et al. (1999), the macroeconomic models are "designed to help us understand the main pressures on the volume of crime at a macro-level and to give an indication of how the crime level might change in the future, reflecting these underlying pressures".

Macro level or time-series approach to crime data is not a new phenomenon and there exist a vast literature using the UK and the US data. Existing literature on macroeconomic aspects of crime often applies univariate error correction models (ECM), vector autoregression (VAR) and vector error correction (VECM) models. Some of the time-series references using the UK data are Pyle and Deadman (1994), Dhiri et al. (1999), Hale and Sabbagh (1991); examples of papers using the US data are Greenberg (2001), O'Brien (1999), Corman et al. (1987), Cappell and Sykes (1991), Witt and Dryden Witte (1998), and Saridakis (2004). In Switzerland, examples of papers which analyse crime rates using a time-series approach are papers by Funk and Kugler (2003a) and Funk and Kugler (2003b).[2]

Estimation results in these papers are based on the cointegration modelling approach, which assumes a stable long-run relationship between crime and underlying explanatory variables.[3] This relationship is allowed to be in a disequilibrium in the short run, error-correcting to its long-run equilibrium. The question is, why should there be a long-run, stable equilibrium for crime? Deadman (2003) shows that forecasts of recorded burglary and theft for England and Wales to the year 2001, based on the estimated cointegration relationship, seriously overestimated its realisations and created unnecessary unrest in the eyes of the public about the expected crime trends.[4] Additionally, it has been seen in the literature that the relevant crime economics data, especially demographic data, are often integrated of second order, meaning that they need to be differenced twice in order to become stationary. This poses a great difficulty when trying to specify a satisfactory cointegrated model for crime variables, as has been encountered by Boswijk et al. (2002) and Vujić (2003).

This chapter proposes a new time-series approach to studying dynamics of crime. In particular, we approach the macroeconomic relationship between crime and economic variables using the unobserved components time series models. We see this as our first contribution to the existing crime-economics literature. The unobserved components time series, also known as structural time series (STS) models, have not extensively been used in the field of economics of crime.[5] The analysis is based on state space techniques

[2]For the Netherlands, the often cited paper that studies trends in recorded crime is by Beki et al. (1999). The authors estimate a relationship between recorded crime and demographic, economic and policy changes in the Netherlands in the period 1950-1993. Another paper by Moolenaar et al. (2002), estimates the demand for sanction capacity for juvenile delinquents in the Netherlands using aggregate data from 1965 to 1998. A paper by Boswijk et al. (2002), uses data from 1955 to 2000 to relate trends in crime in the Netherlands to demographic, socio-economics and law enforcement performance indicators. Using Monte Carlo simulations, this paper also forecasts crime trends for the period 1995-2000.

[3]An exception is a paper by Cappell and Sykes (1991), who base their results on the ARIMA time-series approach.

[4]In the words of Deadman, "...the models that incorporated an error correction component gave substantial over-predictions and a completely wrong pattern of outcomes." And further, "perhaps the most important implication of these results is to throw doubt on the concept of a long-run equilibrium for crime." Britt (2001) writes that the cointegration approach "does not allow for time-varying parameters, which could address concerns about ahistorical statistical models."

[5]To our knowledge, the structural time series methodology applied to crime data is carried out by Harvey and Fernandez (1989) and Atkinson et al. (1997), who look at the number of outliers and breaks in the monthly number of purse (handbag) snatches in Hyde Park in Chicago. Koopman et al. (2008) model recidivism behaviour of juveniles from a Dutch Judicial Juvenile Institution, using a non-Gaussian

and, in particular, the Kalman filter.[6] Structural time series models typically consist of interpretable components such as trend, seasonal, cycle, and irregular component. Since each one of the components can be modelled explicitly, they allow for time-varying parameters in the observation regression. As it is argued by Beveridge and Nelson (1981), Granger and Newbold (1976), Clark (1987), and Harvey (1989), many macroeconomic time series can be decomposed in a persistent component (nonstationary trend) and a transitory component (stationary cycle). If we agree that the series of recorded burglary and theft data show similar trendy behaviour like the purely economic series such as GDP or unemployment rates, then we have a good starting point for analysing economic aspects of crime using unobserved components models. The similarity in general patterns over time between crime and economic series make us believe that crime, like the economy, may be subject to different short run and long run factors. Additionally, we believe that economic factors are one of the major forces that can explain changes in aggregate crime. Considering that the business cycle has a pervasive effect on the structure of economic opportunity and hence on individual behaviour, it would be surprising if crime rates were immune to general business conditions.

Our second contribution is in explicit examination of the cyclical behaviour in the crime macroeconomic data, where we model the business cycle as a stationary autoregressive moving average (ARMA) process. "The relationship between general business cycle condition and crime has been a popular subject of study for empirical criminologists since the mid-nineteenth century" (Cook and Zarkin, 1985), particularly in England and Wales (see for example Swaine-Thomas, 1927 and Wolpin, 1978). In the previous literature, Cantor and Land (1985) model changes in the business cycle as year-to-year differences in crime rates, which is too short of a period to capture changes in crime caused by the business cycle. Cook and Zarkin (1985) measure effects of the business cycle as deviations from the secular trend in the data, by dividing crime and unemployment rates by the 3− or 5−year moving averages. This a better measure of business cycle, because business cycle is a longer than one year phenomenon. Corman et al. (1987) apply a VAR approach in order to study crime, deterrence and the business cycle in New York city. Paternoster and Bushway (2001) question whether VAR approach allows us to capture well the effect of business cycle changes on crime rates. In a more recent paper, Arvanites and Defina (2006) measure the effect of a business cycle by employing inflation-adjusted (real) gross state product (GSP). Other papers that examine the effect of changes in the economy on changes in crime data relate analysed crime series to measures of unemployment, consumption, GDP, and/or inflation and conclude that criminal offences are sensitive to the cycle in economic activity (for an extensive literature overview on crime and economic incentives, see Section 3.4). Most papers find that property crimes have a counter-cyclical pattern, falling and rising along with the ups and downs of the cycle in economic activity, while violent crimes are not so much business cycle driven (except maybe for robbery, which has a clear for-profit motivation). We think that relating levels of unemployment, consumption, GDP, and/or inflation as measures of economic activity to levels of crime is

STS model. Finally, Vujić and Koopman (2008) (chapter 2 of this dissertation) investigate the impact of sentence reform in Virginia (US) in the 1990s on the recorded crime rates.

[6] In the crime literature, suggestions of the application of the Kalman filter have been made by Britt (2001).

not sufficient to conclude that business cycle fluctuations have been captured (for en excellent discussion, see paper by Paternoster and Bushway, 2001). Hence, the question we pose is: Is there a cycle in crime? If crime series are sensitive to the cyclical behaviour of the economy, we are then particularly interested in testing whether a cycle in crime data only coincides with the business cycle pattern or has crime its own cyclical behaviour?

In order to study economic-induced behaviour in recorded burglary and theft data in England and Wales in the period 1955 to 2001, we adopt both univariate and multivariate time series frameworks within the unobserved components models. We interpret 'multivariate analysis' in two ways. On one hand, we look for a multivariate explanation of crime movements that typically includes several measures of economic conditions, demographic and deterrence variables. Economic theory provides us with sufficient justification to explore crime data in a multivariate framework and this is a standard approach in empirical studies of crime and economics conditions. On the other hand, burglary, theft, GDP, unemployment, and police variables have been made endogenous, which is a novelty of our approach.[7] We see estimation of the simultaneous (multivariate) model of crime and economic variables as the third contribution of this chapter.

The fourth contribution of this chapter lies in empirically establishing the opportunity and motivation effects of criminal behaviour, introduced in the seminal paper by Cantor and Land (1985). They discuss two channels through which business cycles can affect the rate of criminal activity. In the short-run, improving economy means more suitable and more attractive targets for criminals, which translates into pro-cyclical crime behaviour (rising economy and rising crime). This is referred to as the opportunity effect. In the long-run, changing economic conditions affect criminal behaviour through social strain and social control, which translates into counter-cyclical crime behaviour (rising economy and decreasing crime rates). This is referred to as the motivational effect. Although many papers have empirically established the long-run counter-cyclical pattern of property crime (see Section 3.4), to our knowledge, no paper has managed to empirically establish the criminal opportunity effect. This chapter sheds light on the presence of criminal opportunity and motivational effects.

The empirical findings of our study can be summarized as follows. We have found strong evidence for cyclical behaviour in burglary and theft rates. Further, it is found that the cyclical dynamics in burglary and theft rates can be fully accounted for by cycles of economic activity. We therefore conclude that we have found no evidence of a separate crime cycle, it can be fully explained by economic cyclical behaviour. The presented analysis has accounted for various other variables that can explain variations in burglary. The explanatory variables affect more the short-term dynamics than the long-term dynamics. Another interesting finding of this chapter is that the model has detected the motivational and opportunity effects of burglary and theft in relation to economic

[7]Simultaneous dependence of crime and economic conditions has been previously explored in a theoretical paper by Burdett et al. (2003), where crime, inequality, and unemployment are made endogenous using an equilibrium search model. Further, Corman et al. (1987) apply a VAR approach in order to study crime, deterrence and the business cycle in New York city. Additionally, Thornberry and Christenson (1984) show that unemployment and crime mutually influence one another over the individual's life span, such that unemployment has significant instantaneous effects on crime, while crime primarily has significant lagged effects on unemployment.

conditions.

The chapter is organised as follows. Section 3.2 describes the data, discusses the unobserved components modelling approach and presents the univariate decomposition results. Section 3.3 discusses the multivariate unobserved components modelling approach and presents empirical findings of the multivariate decomposition results. Section 3.4 relates our findings with the existing empirical literature on crime. Section 3.5 concludes.

3.2 Decomposition of the economic and crime series

3.2.1 Data and methodology

To gain insights into the properties of the yearly time series for England and Wales, we decompose the two economic time series of real gross domestic product (GDP) and number of unemployment (Unemp) together with the two crime time series of burglary per capita (Burg) and theft per capita (Thft) for the years 1955 to 2001[8]. These four time series (in logs) are selected since we are interested in the interactions and relationships between crime and economic activity. Various explanatory variables explain crime such as criminal justice variables (conviction rate, sentence length, and imprisonment), demographic variables (male, youths aged 15-24 years), the effect of a change in recording practice (a dummy variable accounting for the introduction of the Theft Act in 1968),[9] and the effect of number of police officers. Conviction rate is a measure of sentence probability, sentence length and imprisonment rate are measures of severity of punishment, while the number of police officers indicates the protection provided by the state. We will consider these variables in our univariate as well as in our multivariate model analyses. The explanatory variables will be treated as exogenous except for the number of police officers. Since police numbers are the result of economic conditions and crime figures, we will treat this variable as endogenous in a later part of this study. The endogeneity of police hiring seems intuitive: when crime rates go up, the public policy response is to hire more police.

The four time series of our primary concern are presented in Figure 3.1. Although the series have different characteristics, the dynamic features of the time series are somewhat similar. While the four time series have positive growths, in the early 1990s the numbers of unemployment, burglary and theft start to decline. Cyclical fluctuations from the trends are visible in all four series. The cycle of economic activity reflected in the unemployment figures appears quite closely matched to alterations in the level of recorded property crime. We see relatively large rises and falls in recorded crime from about 1980 to 1992, and the sustained fall from this date onward, that match the cyclical movement in unemployment rates over similar periods. To analyse such features in the time series in more detail, we consider the unobserved components time series model as given by

$$y_t = \mu_t + \psi_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}\mathrm{ID}(0, \sigma_\varepsilon^2), \quad t = 1, \ldots, n, \tag{3.1}$$

[8]The data are obtained by courtesy of Derek Deadman from the University of Leicester. The formal descriptions of the time series are given in Appendix 3.A

[9]See `http://www.lawteacher.net/Criminal/Property%20Offences/` for more on Theft Act 1968.

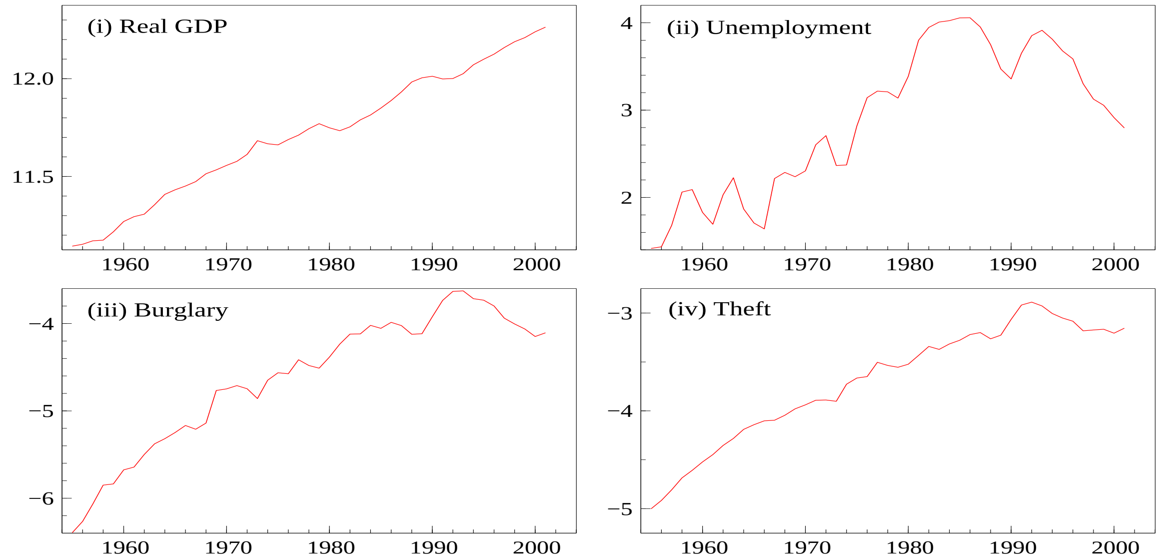

Figure 3.1: Yearly time series for England and Wales between 1955 and 2001, in logs: (i) real gross domestic product (GDP); (ii) number of unemployment per capita; (iii) number of reported burglaries per capita; (iv) number of reported thefts per capita.

where y_t represents the observed time series, trend μ_t and cycle ψ_t are stochastic functions of time while the irregular ε_t is treated as Gaussian white noise.

The trend function can be based on a time-varying constant μ_t for which the time-variation can be modelled as a random walk. To incorporate a growth term in the trend, a time-varying drift β_t can be considered. The resulting model is known as the local linear trend model and is given by

$$\begin{aligned} \mu_{t+1} &= \mu_t + \beta_t + \eta_t, & \eta_t &\sim \mathcal{N}\mathrm{ID}(0, \sigma^2_\eta), \\ \beta_{t+1} &= \beta_t + \zeta_t, & \zeta_t &\sim \mathcal{N}\mathrm{ID}(0, \sigma^2_\zeta), \end{aligned} \tag{3.2}$$

where the disturbances η_t and ζ_t are mutually independent from each other. The fixed trend ($\sigma^2_\eta = \sigma^2_\zeta = 0$), the random walk ($\beta_1 = 0$ and $\sigma^2_\zeta = 0$), the random walk plus drift ($\sigma^2_\zeta = 0$) and the trend from the Hodrick-Prescott filter ($\sigma^2_\eta = 0$ and $\sigma_\varepsilon \ / \ \sigma_\zeta = 14400$) are special cases of (3.2). The estimation of the variances σ^2_ε, σ^2_η and σ^2_ζ by the method of maximum likelihood will determine the dynamic characteristics of the trend specification.

The cycle component can be modelled as a stationary autoregressive moving average (ARMA) process. In the macroeconomics literature the cycle component is sometimes represented by an autoregressive process of order 2, that is an AR(2), see Clark (1987). To ensure that the implied autocorrelation function of the ARMA process exhibit cyclical properties, the roots of the autoregressive polynomial requires to be within the complex range. This can be enforced by considering the bivariate process

$$\begin{aligned} \psi_{t+1} &= \phi\cos(\lambda)\psi_t + \phi\sin(\lambda)\dot{\psi}_t + \kappa_t, & \kappa_t &\sim \mathcal{N}\mathrm{ID}(0, \sigma^2_\kappa), \\ \dot{\psi}_{t+1} &= \phi\cos(\lambda)\dot{\psi}_t - \phi\sin(\lambda)\psi_t + \dot{\kappa}_t, & \dot{\kappa}_t &\sim \mathcal{N}\mathrm{ID}(0, \sigma^2_\kappa), \end{aligned} \tag{3.3}$$

where $|\phi| < 1$ is the damping factor and λ is the cycle frequency. The damping term ϕ

ensures that the stochastic process (3.3) is a stationary ARMA(2,1) process. The period of the cycle is given by $2\pi/\lambda$ and frequency λ indicates that the peak of the frequency spectrum of the ARMA process is at λ while the sharpness of the peak is determined by ρ and $\sigma_\kappa/\sigma_\varepsilon$. The special case of $\rho = 1$ and $\sigma_\kappa^2 = 0$ reduces the stochastic cycle component to a deterministic cycle or a constant cosine wave with frequency λ. More details on the time- and frequency-domain properties of different trend and cycle specifications are given in Harvey (1989). In order to provide a more accurate description of the stationary properties in the time series of interest, we consider a sum of two cycle processes in the model

$$\psi_t = \psi_t^{(1)} + \psi_t^{(2)}, \tag{3.4}$$

where each $\psi_t^{(j)}$ is formulated as in (3.3) for $j = 1, 2$, but each cycle process may have different values for ϕ, λ and σ_κ^2. The interpretation of each cycle mainly depends on the value of its period $2\pi/\lambda$. The first cycle can represent the long-term cycle and the second cycle can refer to the short-term cycle. Finally we note that all disturbance terms in the model are mutually independent of each other, contemporaneously and at all lags.

For a given parameter vector θ that consists of variances, damping factors ϕ and frequencies λ, the model can be represented as a linear Gaussian state space model. The unobservable processes μ_t and ψ_t and their auxiliary components (β_t and $\dot{\psi}_t$) are placed in a so-called state vector. The one-step ahead predictor of the state vector at time t, given the observed values for $y_1, \ldots, y_{t-1}$, is evaluated recursively by the Kalman filter. The prediction error and its variance are also evaluated by the Kalman filter. As a result, the Kalman filter constructs the loglikelihood function via the prediction error decomposition. The maximum likelihood estimate of θ is then obtained by the numerical maximisation of the loglikelihood function with respect to θ. Based on the maximum likelihood estimate of θ, the trend-cycle decomposition can be carried out by applying a so-called smoothing algorithms to estimate the unobserved state vector conditional on all observations $y_1, \ldots, y_n$. A more detailed account of state space methods for the analysis of time series can be found, for example, in Durbin and Koopman (2001).

3.2.2 Trend-cycle decompositions for economic and crime series

The estimation results of the univariate decomposition models are reported in Model 1 of Table 3.1. We first report the estimated standard deviations and other parameters of the individual components. They provide information about the dynamic characteristics of the estimated components. The smoothest trend component is obtained for GDP (the smallest estimated standard deviation of the trend component). The medium cycles have periods in the range of 11-12 years while the short cycle periods are close to 5 years (except for unemployment where the short cycle lasts 6.5 years). The persistences of the cycles are all estimated as being close to 0.9 except for the short cycles in the crime series which are closer to 0.8.

When the Gaussian UC model is correctly specified, the standardised one-step ahead prediction errors produced by the Kalman filter should be a serially uncorrelated sequence of standard normal random numbers. The reported residual diagnostics are for these prediction errors and indicate whether the assumptions of normality (Bowman-Shenton

χ^2 N statistic) and no serial correlation (Box-Ljung portmanteau χ^2 Q statistic) are valid. We conclude that these statistics are satisfactory except for a somewhat large normality test statistic for real GDP. Furthermore, the dynamic behaviour in Burglary is not explained perfectly in the univariate analysis since the Box-Ljung statistic has the relatively high value of 16.5.

In general, the trend plus two cycle components model provides an effective and parsimonious description of the dynamic characteristics in the time series. It should however be emphasized that the cycles play an important role in both the economic and crime series. When the cycle components are removed from the unobserved components model, the goodness-of-fit R_D^2 becomes much lower and is sometimes even negative. A negative R_D^2 implies that a random walk model provides a better description of the data than the model under investigation. Furthermore, the Box-Ljung statistics are all much higher. Further discussion of the importance of stochastic cycles in the time series is given below.

The univariate decompositions are graphically presented in Figure 3.2. The salient features of the time series become apparent in these trend-cycle decompositions. The trends show positive growth rates for all series. Only from the mid–1980s the growth rate of unemployment becomes negative and from the mid–1990s the growth rates of burglary and theft become negative. The cycles capture mainly the different deviations from the trend. It is interesting to detect the negative relation between the medium cycles of unemployment and real GDP. This is a well-known macro-economic phenomenon – a growing economy leads to smaller numbers of workers unemployed. The cases for burglary and theft are less known and are the main focus of this empirical study. We therefore present in Figure 3.3 two plots with the estimated short and medium cycles for the four time series. It shows that both the short and the medium crime cycles are mostly positively correlated with the unemployment cycle and therefore also negatively correlated with the real GDP cycle. The actual sample-correlations of the estimated cycles components are given by

	GDP	Unempl	Burgl	Theft
GDP	1	-0.74	-0.76	-0.44
Unempl	-0.53	1	0.73	0.24
Burgl	-0.54	0.12	1	0.43
Theft	-0.58	0.20	0.73	1

where the upper triangular part presents the correlation for the medium cycles while the lower triangular part is for the short cycles. This preliminary analysis shows that cyclical fluctuations in the numbers of unemployed are positively related to those of burglary and theft. This is a preliminary indication that at the throughs of the cycles in economic activity, unemployment rates rise together with the rates of burglary and theft convictions. Whether such preliminary conclusions hold in a more elaborate and simultaneous model setting is discussed in Section 3.3.

3.2.3 Regression effects

Regression effects can be added to the unobserved components time series model in a standard fashion, see Harvey (1989). In the context of this study, we can represent the

Table 3.1: Maximum likelihood estimation of univariate models for real GDP, Unemployment, Burglary and Theft (sample 1955-2001, without and with regression effects)

		Model 1				Model 2				Model 3	
Parameters		GDP	Unemp	Burgl	Theft	GDP	Unemp	Burgl	Theft	Burgl	Theft
trend	stdev ×100	0.25	3.36	1.54	1.09	0.28	9.38	3.45	0.51	0.98	0.65
cyc-m	stdev ×100	0.92	7.75	1.09	3.25	0.90	0	0	0	4.07	3.72
	period	12.1	10.7	11.1	12.7	12.3	12.0	8.93	15.5	15.4	15.9
	rho	0.89	0.88	0.99	0.85	0.92	1	1	1	0.95	0.79
cyc-s	stdev ×100	0.50	3.89	3.38	1.55	0.30	5.78	2.59	2.56	2.04	0
	period	5.08	4.65	6.49	5.22	4.79	4.64	4.80	6.46	5.33	5.07
	rho	0.87	0.95	0.83	0.78	0.92	0.91	0.74	0.70	0.81	1
irreg	stdev ×100	0	0	3.99	0	0	0	0	0	0	0
Regression						GDP	Unemp	Burgl	Theft	Burgl	Theft
Act 1968	×100					**−3.66**	24.6	**36.6**	4.52	**31.1**	6.99
Young M	×100					−8.14	18.2	118	**88.7**	**176**	33.4
Sentenc B	×100					**−9.52**	15.7	**36.9**	**19.7**	**29.9**	−−
Impris B	×100					−1.78	**74.9**	3.68	−9.54	−20.5	−−
Convict B	×100					−2.83	−5.19	**−59.4**	**−29.8**	**−61.1**	−−
Sentenc T	×100					**14.53**	−7.64	−40.6	**−38.2**	−−	**−32.4**
Impris T	×100					3.29	**−66.2**	**−30.5**	−16.9	−−	**−20.8**
Convict T	×100					4.02	7.78	7.14	−5.62	−−	**−34.4**
Diagnostics		GDP	Unemp	Burgl	Theft	GDP	Unemp	Burgl	Theft	Burgl	Theft
N	$\chi^2(2)$	9.44	1.73	3.50	3.70	1.60	3.28	0.49	0.53	0.02	0.81
$Q(13)$	$\chi^2(6)$	5.72	3.89	16.5	6.75	4.23	8.84	6.80	1.96	7.30	9.96
R^2_D		0.16	0.43	0.22	0.18	0.42	0.43	0.66	0.68	0.61	0.48

Note: All models includes trend component μ_t as in (3.2), medium $\psi_t^{(1)}$ and short $\psi_t^{(2)}$ cycles as in (3.3) and irregular component ε_t as in (3.1). The standard deviations are associated with the disturbances of the different components. For regression effects, values in bold are significant at a 95% significance level. Diagnostics reported are the Bowman-Shanton normality test N, the Ljung-Box test for serial correlation $Q(13)$ and goodness-of-fit criterion R^2_D (compared with fit of random walk plus drift model). Explanatory variables related to burglary and theft are referred to with B and T, respectively. The number of observations is 47.

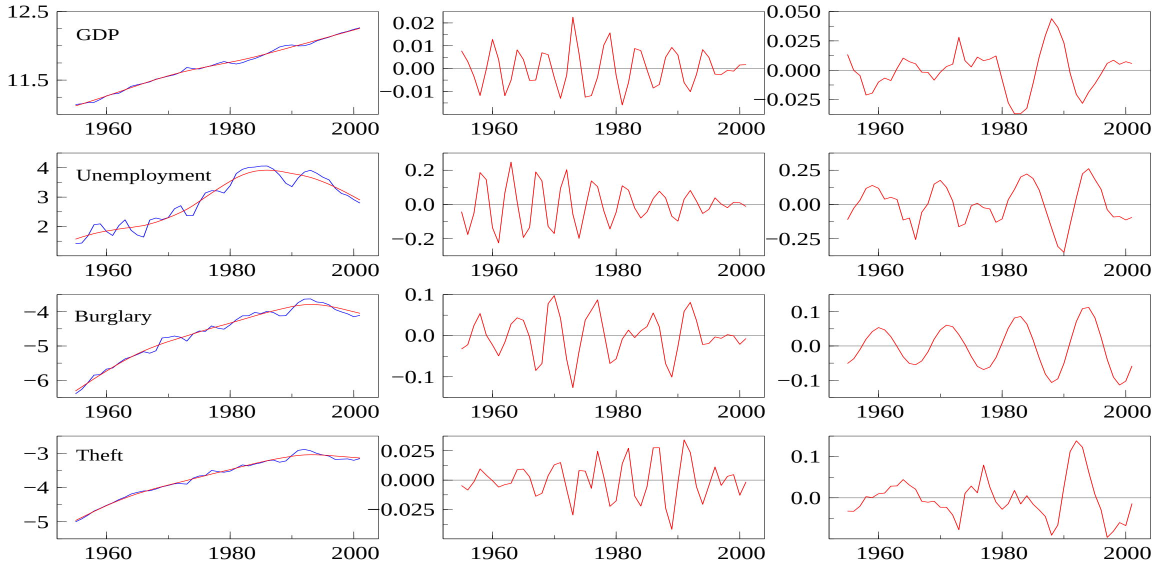

Figure 3.2: Trend-cycle decompositions are based on the univariate unobserved components model (3.1), (3.2) and (3.4). The estimated decompositions are presented for (row-wise) real GDP, Unemployment, Burglary and Theft. The estimated components are for (column-wise) trend, short cycle and medium cycle. The estimation is carried our for the sample 1955-2001.

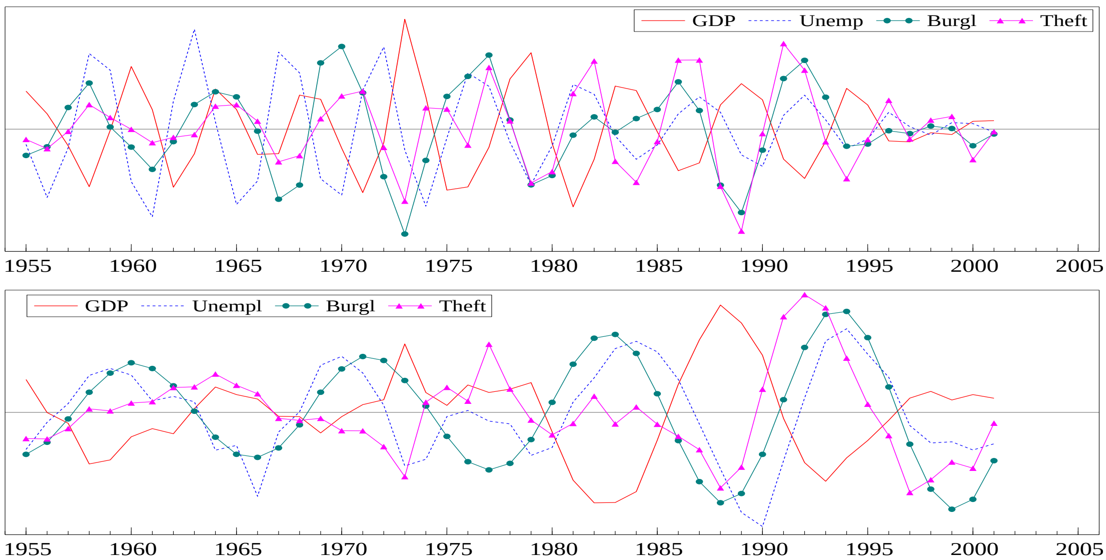

Figure 3.3: Estimated cycles (scale-adjusted by range) are based on the univariate unobserved components model (3.1), (3.2) and (3.4). The first plot presents the short cycles for real GDP, Unemployment, Burglary and Theft. The second plot present the medium cycles for the same series. The estimation is carried our for the sample 1955-2001.

extended model by a regression model with a time-varying intercept and autoregressive moving average ($ARMA$) disturbances. In particular, we consider the regression model

$$y_t = \mu_t + x_t'\delta + u_t, \qquad u_t \sim ARMA(p,q), \tag{3.5}$$

where μ_t is modelled as (3.2), x_t is the $k \times 1$ vector of exogenous explanatory variables, δ is the $k \times 1$ vector of regression coefficients and disturbance u_t is modelled as $ARMA(p,q)$ with p and q representing the lag orders for the autoregressive and moving average parts, respectively. In the case of the unobserved components model with two stationary cycle processes, the properties of the ARMA disturbances are implied by the cycle and irregular components. It should be noted that sums of ARMA components can always be represented by a single ARMA process and that cycle and irregular components are constrained ARMA processes. The estimation of δ is based on the methods of generalised least squares or maximum likelihood and is carried out simultaneously with the maximum likelihood estimation of the parameters associated with the components as described in Section 3.2.1. We refer to Durbin and Koopman (2001) for a discussion of different estimation methods when regression effects are present in an UC model.

A set of eight explanatory variables for the analysis of the crime series is suggested in Section 3.2.1. These are the dummy for the 1968 Burglary Act, the number of males aged between 15-24 years, the average length of sentences for Burglary and Theft, the percentage of convictions that are imprisoned for Burglary and Theft and the number of convictions for Burglary and Theft. It follows from these definitions that the explanatory variables are expected to be highly correlated with each other. For example, the correlation between the conviction rate of Burglary and Theft is 0.94 while for imprisonment it is 0.84, but for sentence length it is only 0.24. A straightforward principal components analysis for the eight explanatory variables indicates that the first 2 eigen-vectors represent 76.0 percent of the variation while the first 5 eigen-vectors represent 98.8 percent. Therefore a parsimonious model will be preferred in the empirical study. However, we first present some standard regression results based on the dynamic regression model (3.5) where all eight explanatory variables are included.

The four univariate regression models are estimated by maximum likelihood separately with the inclusion of all explanatory variables. A priori, the explanatory variables should not affect GDP and Unemployment since the explanatory variables are solely related to crime phenomena. Also, the explanatory variables related to Theft should improve the fit of the Theft crime series more than the fit of Burglary and vice versa. Since the purpose of this study is to develop a simultaneous model for economic and crime series, we initially carry out the regression analyses with all eight explanatory variables for all four dependent series. It also provides a robustness check to model misspecification since we can clearly select the variables that should be significant and those that should not. The regression results are presented in Model 2 of Table 3.1 together with the estimated coefficients associated with the unobserved components and the diagnostic statistics.

It is not surprising that the fit for all four UC models has improved as a result of including the regression effects. In comparison with the R_D^2 value reported in Model 1, the R_D^2 value in Model 2 triples for the two crime series while for GDP it doubles and for unemployment it remains the same. This confirms our initial suggestion that the explanatory variables should overall have a bigger impact on the crime series. A closer inspection of the results in Model 2 further shows that the regression effects have mainly an impact on the shorter and medium dynamics (the cycle and irregular components) rather than the long term trend component. In the Burglary regression, the irregular disappears and for all series but GDP the shorter cycle reduces to a deterministic cycle ($\rho = 1$ and $\sigma_\kappa = 0$). The regression effects are therefore able to explain the noise that is unaccounted for in a pure time series model. The noise is related to high spectral frequency components such as the irregular and short cycle components. From an empirical standpoint, changes in measures to combat crime have primarily a direct effect on the short-term behaviour of the analysed crime series. The regression results in Model 2 confirm this.

Since the regressors are related to crime phenomena, unemployment and real GDP should not be affected by these explanatory variables. Some of the regressions results are therefore regarded as spurious. For example, the Theft Act of 1968 should not affect GDP but it is estimated as a significant effect in this univariate analysis. The regression results can be summarized further by observing that 13 significant regression effects are found in the four regressions, 5 out of these 13 should not be significant (they occur in GDP and Unemployment). However, the explanatory variables corresponding to these 5 significant effects are also significant in the crime series. This may indicate that both GDP and Unemployment have an impact on the crime series. The simultaneous modelling results to be presented in Section 3.3 will clarify these issues. Furthermore, 8 significant coefficients are found in the two crime equations from which 2 of them have incorrect signs and both are for the average length of sentences for Burglary. This may be due to the high correlation between the explanatory variables. The other significant coefficients have correct signs. For example, more imprisonment and higher conviction rates should have a negative effect on burglary rates. Although the overall regression results are not perfect, sufficient evidence is given that a simultaneous model for the economic and crime series should include regression effects. Furthermore these results illustrate that obtaining theoretically consistent results from a univariate analysis is not straightforward when the sample length is short. This is yet another motivation to develop a simultaneous model in the next section.

Finally, the explanatory variables for average sentence length, numbers of imprisonment and conviction rate are separately available for Burglary and Theft categories. Burglary should mainly be affected by the explanatory variables related to Burglary and for Theft by those related to Theft. Also it is reported earlier that the explanatory variables are highly correlated amongst themselves and a robustness check of the earlier results can be useful in this respect. Therefore Model 3 of Table 3.1 presents the results for Burglary and Theft where only those regression variables are included that relate to the corresponding category of crime. The estimated dynamics are closer to those of Model 1 and therefore less affected by the inclusion of the regression effects. A movement from the long-term innovations (trend component) to the medium-cycle innovations can be observed by comparing the estimated standard deviations of the trend and the medium cycle in Models 1 and 3. For Burglary, it is interesting that the regression effects still replace the unaccounted noise (the irregular component) that is present in the pure time series model of Model 1. The regression results for Theft become less convincing when the regression restrictions are imposed. The fit and the diagnostics are weaker compared to those of Model 2. However, the significantly estimated regression coefficients in Model 3 have correct signs for Theft. The fit of the Burglary equation is much less affected by the exclusion of the Theft related explanatory variables. The significantly estimated regression coefficient for average sentence length remains to have the wrong sign for Burglary.

3.3 Multivariate trend-cycle decompositions

3.3.1 The multivariate unobserved components model

A simultaneous analysis of a set of multiple time series collected in the $N \times 1$ vector y_t can be based on the decomposition model (3.1) with the scalar components replaced by vector components. We define the multivariate unobserved components model by

$$y_t = \mu_t + \psi_t + X_t\delta + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}\text{ID}(0, \Sigma_\varepsilon), \qquad t = 1, \ldots, n, \tag{3.6}$$

where trend μ_t and cycle ψ_t are stochastic $N \times 1$ vectors and X_t is an $N \times k$ matrix of exogenous variables. The regression coefficient vector δ is fixed and unknown. The $N \times 1$ irregular vector ε_t is generated by the multivariate normal distribution with zero mean and variance matrix Σ_ε. The dynamic model specifications (3.2) for trend μ_t and (3.3) for cycle ψ_t remain as in the univariate case but the scalar process becomes a vector process. The disturbances in these stochastic specifications also become vectors and are assumed to come from multivariate normal distributions. Specifically, we have

$$\eta_t \sim \mathcal{N}\text{ID}(0, \Sigma_\eta), \quad \zeta_t \sim \mathcal{N}\text{ID}(0, \Sigma_\zeta), \quad \kappa_t \sim \mathcal{N}\text{ID}(0, \Sigma_\kappa), \quad \dot{\kappa}_t \sim \mathcal{N}\text{ID}(0, \Sigma_\kappa), \quad t = 1, \ldots, n.$$

The cycle specification (3.3) has become a vector equation but the discounting factor ϕ and cycle frequency λ are common to all elements of ψ_t. These coefficients are therefore kept as scalars. Harvey and Koopman (1997) define this specification as the similar cycle model. To incorporate cycles with different frequencies in the model, different similar cycle components can be included.

This straightforward multivariate extension of the trend-cycle decomposition model (3.1) is referred to as the seemingly unrelated time series equation (SUTSE) model. The individual slopes in vector β_t are only related through the correlations between the individual disturbances in vector ζ_t as implied by the variance matrix Σ_ζ. The same principle applies to the slope and cycle vector components in the model. The disturbance variance matrices therefore play an important role. In particular, the rank of the variance matrix is of interest. For example, in case Σ_ζ has full rank, all trend disturbances in ζ_t have their own unique source of variation but may be correlated between each other. In case Σ_ζ has lower rank, the individual trend disturbances in ζ_t are generated by a smaller set of independent disturbances. This follows straightforwardly since any variance matrix can be expressed via the Choleski decomposition, that is

$$\Sigma_\zeta = A_\zeta D_\zeta A_\zeta', \tag{3.7}$$

where A_ζ is a lower unity triangular $N \times r_\zeta$ matrix and D_ζ is a diagonal $r_\zeta \times r_\zeta$ matrix with the rank of Σ_ζ given by r_ζ. In a strict sense, we require a full column rank matrix for

A_ζ and positive values on the diagonal of D_ζ for matrix Σ_ζ to have rank r_ζ. Consequently, we have

$$\zeta_t = A_\zeta \zeta_t^*, \qquad \zeta_t^* \sim \mathcal{N}\mathrm{ID}(0, D_\zeta),$$

and

$$\beta_t = \beta^* + A_\zeta \beta_t^*, \qquad \beta_{t+1}^* = \beta_t^* + \zeta_t^*, \tag{3.8}$$

where β^* is a fixed vector with $N - r_\zeta$ non-zero values and r_ζ zero values. In this way, all slope variables in β_t have different initial values. In case $r_\zeta < N$ and as a consequence Σ_ζ has a lower rank, ζ_t and β_t are linear combinations of a smaller set of stochastic processes. The same arguments apply to other disturbances and components in the model. A lower rank of the variance matrix can be imposed but it can also be the result of estimation. Testing procedures for common trends and cycles are developed recently by Nyblom and Harvey (2001).

By adopting the notation implied by (3.7) and (3.8), we can reformulate model (3.6) in terms of common factors. Such a decomposition model with trend, two cycles and irregular components together with regression effects is given by

$$y_t = \mu^* + A_\eta \mu_t^* + A_\kappa^{(1)} \psi_t^{*(1)} + A_\kappa^{(2)} \psi_t^{*(2)} + X_t \delta + A_\varepsilon \varepsilon_t^*, \tag{3.9}$$

where factor loading matrices A_η, A_ζ, $A_\kappa^{(1)}$, $A_\kappa^{(2)}$ and A_ε are lower unity triangular $N \times r$ matrices (with $r \leq N$ varying for each loading matrix) and the common components μ_t^*, β_t^*, $\psi_t^{*(1)}$, $\psi_t^{*(2)}$ and ε_t^* are associated with disturbances that have diagonal variance matrices. Common components are of interest for studying the dynamic structures and interactions within a set of time series. For example, common trends imply that economic time series are cointegrated, see the discussion by Stock and Watson (1988). The estimation of the unobserved components trend and cycle and the maximum likelihood estimation of unknown coefficients is similar as described in the previous section since state space methods are applicable to multivariate models too. However, the computational effort becomes more involved when the dimension of y_t increases. It may also be numerically more difficult to find the maximum of the likelihood function with respect to a large dimensional parameter space.

3.3.2 Multivariate decomposition results

We adopt a multivariate time series approach to study the linkages between crime and economics since relevant variables are mutually causal in a dynamic way. It is hard to identify such linkages using standard regression techniques, see Fisher and Nagin (1978). Additionally, multicollinearity among the independent variables makes it difficult to estimate the relative contributions of deterrence, demographic and economic variables, see Freeman (1983). Eide (2000) argues that the regression approach is not appropriate in the presence of two-way causation. Considering our univariate estimation results and these findings from the literature, we have provided some motivation to analyse the macroeconomic and property crime series simultaneously. The multivariate analysis is first considered for the two vectors of endogenous time series given by

$$z_t^{(1)} = (\text{GDP, Unemp, Burglary})'_t, \qquad z_t^{(2)} = (\text{GDP, Unemp, Theft})'_t,$$

where GDP is real gross domestic product and Unemp is unemployment. The analyses for these time series vectors will provide a first insight into the relation between economic and crime series. The multivariate decomposition model has the same structure as the univariate model with a smooth trend (long term effects), two cycles (short and medium term effects) and irregular (instantaneous effects). The cycle dynamics consist of two cycles. The periods and the persistence of the two cycles are estimated simultaneously, together with the parameters in the variance matrices. Three variants of this model are estimated. The first is the multivariate time series model without explanatory (exogenous) variables, the second is the model with explanatory variables for crime included in all three equations while the third is the model with explanatory variables only in the Burglary or Theft equation. The estimation results for these three models are presented in Table 3.2 for burglary and Table 3.3 for theft. We first discuss the estimation results for the two time series vectors without explanatory variables.

The periods of the two similar cycles are estimated approximately by 5 and 10 years in all three models and these can be regarded as typical economic business cycle frequencies. The amplitude of a cycle indicates its relative importance to explaining the variation in the time series. Although the estimated amplitude of the cycles vary somewhat across the three models, it is clear that the shorter cycle is less important than the longer cycle of 10 years. The shorter 5 years cycle only affects the unemployment series. This can also be observed from the estimated factor loadings of the 5 years cycle component. For example, for the third model, the loadings $(1, -7.2, -2.4)$ are for the first factor with a standard deviation of 0.002. The first factor of the 5 years cycle does therefore vary not so much compared with the second factor that has a standard deviation of 0.03. The loadings for the second factor are given by $(0, 1, -0.015)$ and indicate that this factor is only relevant for unemployment, the second series in y_t.

Therefore we concentrate on the second cycle with a period of approximately 10 years and with a sufficient persistence (discounting factor is estimated between 0.85 and 0.90) in all three specifications. The amplitude of the cycle is largest for unemployment and burglary and therefore the relation between the cycle components of these two series is most interesting. In case of the first model, the cycles of the three series can be described by two cyclical factors and hence the three cycles are based on two underlying common cycles. Given that GDP is only involved with the first common cycle, we can interpret this cycle as the economic business cycle and we confirm its importance for both unemployment and burglary. For the first model, the factor loadings for the business cycle are $(1, -5.2, -4.3)$ and the last two negative coefficients indicate that the economic business cycle of unemployment and burglary are counter-cyclical with GDP. In the economics business cycle literature, it is well-known that unemployment is counter-cyclical with GDP. In the crime economics literature it is also well known that property crimes behave counter-cyclical. Our results confirm these findings.

The second common 10 years cycle factor is concerned with unemployment and burglary only and its standard deviation is five times the standard deviation of the first factor for the first model. Further, this factor is clearly driven by unemployment and may account for the fact that although unemployment is counter-cyclical, its phase may not be the same as for GDP. Unemployment is known to be a lagging indicator of the business cycle and the second common 10 years cycle factor may partly account for this shift. Note that the two common cycles are similar in the sense that they have the same period and persistence. Therefore, the second factor accounts for the cyclical dynamics in unemployment that can not be explained by the business cycle. This correction applies to a lesser extent to burglary as implied by the factor loadings of the second factor, $(0, 1, 0.35)$. Similar results are obtained by the two other models although the correction for burglary by the second factor is less pronounced. For the third model, we have the factor loadings equal to $(0, 1, 0.14)$ from which we conclude that the cycle in burglary is mostly explained by the business cycle.

The trends are modelled by smooth processes where the only variation in the trend is due to changes in the growth which is modelled by a random walk process. The estimation results for the three models all indicate that the long term trend of the three time series is also driven by two common slope factors. The variance matrix of the slope disturbances is of rank 2 as a result. The signs of the factor loadings for the two common slope factors are interesting. The first factor enters the GDP equation only and is therefore interpreted as the long term economic growth. The corresponding factor loadings in the first model is given by $(1, -8.7, 2.7)$. Unemployment growth is strongly negatively related to economic growth and this makes common sense. Burglary growth in the long term is however positively correlated with economic growth and this can be interpreted as the opportunity effect. An increase of wealth in a nation also provides more opportunity for property crime. The equally important second slope factor is only for unemployment and burglary and represents their growth after correction for economic growth. Growth in long term unemployment that is not due to economic growth can be caused by technical development and typically affect less qualified workers. Long-term unemployment may then become a motivational aspect of criminal offences. The factor loadings for the second slope factor in the first model is given by $(0, 1, 0.45)$. The value of 0.45 implies a positive relation of this long-term factor with burglary rates. These long-term effects also apply to models with explanatory variables.

The irregular component represents the instantaneous effect between the three variables. In all models, the irregular for GDP disappears (its variance is estimated as zero). This is typical for the smooth GDP series, irregular noise is not observed as in the case of the univariate results. The irregular series for unemployment and burglary are also relatively small in size but are negatively correlated. This is the short-term substitution effect. It is argued by Cantor and Land (1985) that increased unemployment leads to reduced circulation of people and property. Those who become unemployed have resources (savings, unemployment benefits) which need to be exhausted before they start feeling financially challenged. When more workers become unemployed, the urgency of starting criminal activities is low because of the social benefits they receive. Therefore, crime numbers decrease in the short-run.

3.3.3 Multivariate decomposition results with explanatory variables

The models are estimated for specifications without explanatory variables, with explanatory variables for all three equations and with explanatory variables for only the burglary equation. The explanatory variables are Act 1968 dummy (level intervention), number of young males, sentence length, imprisonment rate, conviction rate and size of police force (see Appendix 3.A for more details).

The general description of the empirical results are given in terms of relationships between the GDP, unemployment and burglary series that are decomposed in trend, cycle and irregular components. The effect of including explanatory variables in the decomposition model does not change these results very much. In case of GDP, number of young males and sentence length are only significant, with t-values of approximately 2.0. The regression effects are most effective for the burglary equation. The goodness-of-fit increases from .42 to .71. Since the explanatory variables are clearly related to crime records, the multivariate model is re-estimated but with the regressors only included in the burglary equation. For the resulting Model 3, the residual diagnostics for all equations are best and therefore we prefer this model. By including the regression variables in the burglary equation, the standard deviation of the slope disturbances is reduced by 17.5%. The standard deviations of the disturbances for the shorter term dynamics are also affected. We therefore do not obtain a clear picture whether the explanatory variables have a more pronounced effect on the short or long term dynamics of recorded burglary rates. All explanatory variables in the burglary equation are significant and of expected sign, except for average sentence length. The magnitude of the estimated coefficients is smaller than in the univariate regression results for burglary.

3.3.4 Model for thefts

We have repeated the calculations for the set of three models but with the burglary time series replaced by the theft time series. These results are reported in Table 3.3. We report these results to show that our findings for burglary are robust and apply to other crime series as well. Theft can be regarded as a crime of a lighter category than burglary. It appears that theft (which we may also refer to as "petty crime") is also affected by macroeconomic conditions. The number of factors and the magnitudes and signs of the factor loadings are estimated similarly as in the burglary case. One exception is the first factor loading for the short cycle of theft on the first factor in Model 3 which is positive while it is negative for the one of burglary. However, the corresponding factor loading for the irregular of theft is more negative compared to the one of burglary. This may suggest that some level of substitution has taken place.

The significantly estimated regression effects for sentence, imprisonment and conviction have the right signs in both Models 2 and 3. In case of Model 2, the explanatory variables are not significant for real GDP and unemployment as it should be the case. The diagnostic statistics for the residuals and the goodness-of-fit results are quite satisfactory. We therefore conclude that the considered trivariate models are adequate model descriptions for the time series of real GDP, unemployment and burglary/theft. As a result, the

Table 3.2: Maximum likelihood estimation results for multivariate decomposition model with real GDP, Unemployment and Burglary (sample 1955-2001, without and with regression effects)

		Model 1			Model 2			Model 3		
Loadings		Fac1	Fac2		Fac1	Fac2		Fac1	Fac2	
trend	GDP	1	0		1	0		1	0	
	Une	−8.7	1		−9.8	1		−10.	1	
	Bur	2.7	.45		4.4	.34		3.3	.32	
cyc-m	GDP	1	0		1	0		1	0	
	Une	−5.2	1		−5.2	1		−5.2	1	
	Bur	−4.3	.35		−3.2	.10		−3.3	.14	
cyc-s	GDP	1	0		1	0		1	0	
	Une	−10.	1		−9.0	1		−7.2	1	
	Bur	.56	−.11		−2.3	−.00		−2.4	−.01	
irreg	GDP	0	0		0	0		0	0	
	Une	1	0		1	0		1	0	
	Bur	−.13	1		−.05	1		−.06	1	
Parameters		GDP	Une	Bur	GDP	Une	Bur	GDP	Une	Bur
trend	stdev ×100	.29	3.3	1.2	.19	3.1	1.2	.22	3.0	.99
cyc-m	stdev ×100	1.1	7.7	5.2	1.2	8.8	3.7	1.2	8.5	3.9
	period	9.8	9.8	9.8	12.	12.	12.	11.	11.	11.
	rho	.86	.86	.86	.89	.89	.89	.85	.85	.85
	ampl ×100	2.6	22.	14.	2.1	20.	7.6	1.7	16.	6.6
cyc-s	stdev ×100	.18	3.4	.34	.17	3.3	.39	.20	3.3	.48
	period	4.6	4.6	4.6	4.6	4.6	4.6	4.7	4.7	4.7
	rho	.96	.96	.96	.96	.96	.96	.95	.95	.95
	ampl ×100	.23	4.0	.32	.19	2.2	.44	.14	2.4	.33
irreg	stdev ×100	0	2.9	3.5	0	1.4	2.6	0	2.2	2.6
Regression effects					GDP	Une	Bur			Bur
Act 68	×100				−1.1	9.2	**28.**			**22.**
Young M	×100				**−21.**	148	**127**			**71.**
Sentence	×100				**−5.7**	22.	**20.**			.77
Impris	×100				−2.5	37.	−5.4			**−12.**
Convict	×100				3.2	3.1	**−67.**			**−59.**
Diagnostic statistics		GDP	Une	Bur	GDP	Une	Bur	GDP	Une	Bur
N	$\chi^2(2)$	3.0	1.5	.80	5.0	.48	.79	3.3	.06	.11
$Q(13)$	$\chi^2(6)$	8.6	8.7	12.	7.3	6.5	7.7	6.2	6.7	5.6
R_D^2		.37	.49	.42	.47	.52	.71	.38	.49	.68

Note: The model includes trend component μ_t in (3.2), medium and short cycles ψ_t in (3.3) and irregular component ε_t as in (3.1). The standard deviations are associated with the disturbances of the different components. For regression effects, values in bold are significant at a 95% significance level. Diagnostics reported are the Bowman-Shanton normality test N, the Ljung-Box test for serial correlation $Q(13)$ and goodness-of-fit criterion R_D^2 (compared with fit of random walk plus drift model). The number of observations is 47.

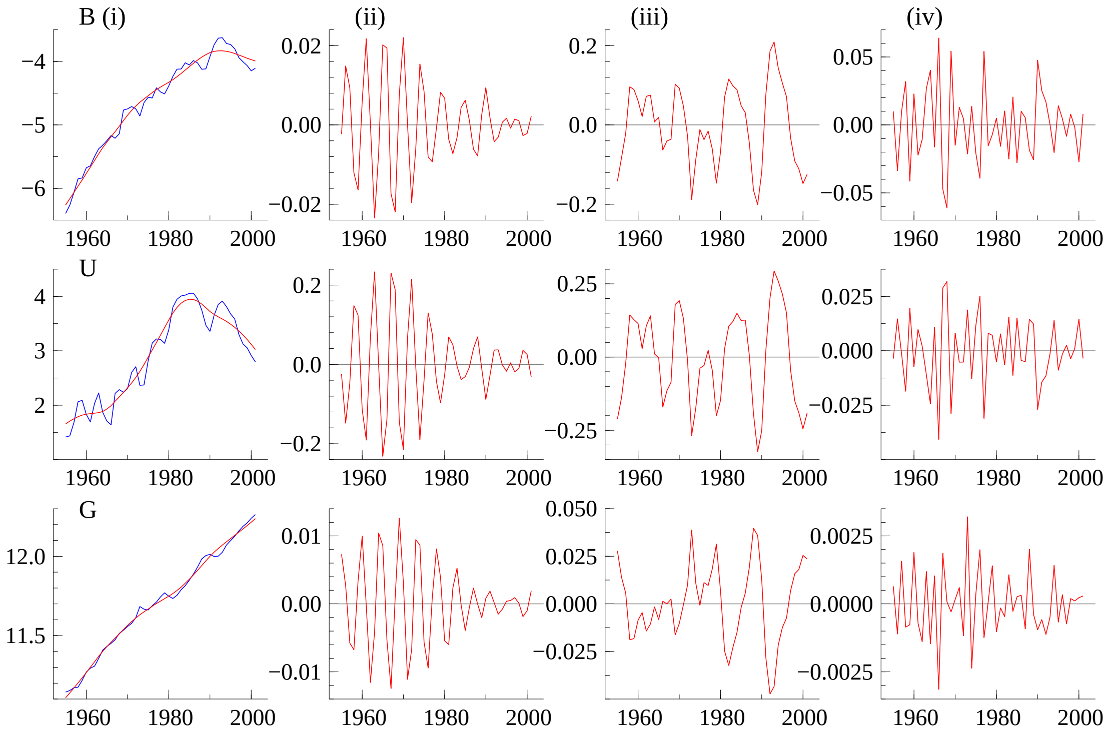

Figure 3.4: Estimated multivariate decomposition for time series – row-wise – B (burglary), U (unemployment) and G (real GDP) with components – column-wise – (i) trend (with observations), (ii) short cycle, (iii) medium cycle and (iv) irregular.

models are appropriate for the investigation of dynamic interrelationships between the three variables.

3.3.5 Multivariate model with economics, crime and police

Changes in the number of police are expected to be positively correlated with changes in the crime rates. The endogeneity of police hiring seems intuitive: when crime rates go up, the public policy response is to hire more police. This has been analysed in papers by Levitt (1997), Draca et al. (2008), Vollaard and Koning (2006). However, Corman et al. (1987) write that "estimation of the supply of crime function has remained a difficulty for economists because of the unknown dynamic interrelationship between the demand for police protection, the production of protection, and the actual supply of crime function."

Funk and Kugler (2003a) analyse the dynamic interactions between crimes, using the VAR approach. Interpreting the impulse response functions, they find that increases in minor crimes dynamically trigger more severe crimes due to the "broken window theory" (if one window is left unrepaired, this will cause other windows to be broken soon) and the life-cycle pattern of criminal careers (an increase in the number of "beginners" who start committing mild crimes, would later increase the number of severe crimes, as they move on along the learning-by-doing approach to crime). If we consider theft a milder crime in comparison to burglary, we would expect a positive causation from theft to burglary.

The estimation results for the multivariate decomposition models for the series of real GDP, unemployment, burglary, theft and police are reported in Tables 3.4 and 3.5. The estimated factor loadings are reported in Table 3.4 and have many common findings with those reported for the trivariate models and which are discussed in previous sections. However, there are some differences. The trend component innovations for burglary and theft are negatively related with the economic growth factor (the first trend factor) while they are positively related with the long-term unemployment factor (the second trend factor) for the three model specifications. The negative relations with the cycle factors of real GDP are also persistent to all model specifications. This finding also applies to police force which implies that with economic upturns the quest for police is less strong since upturns are likely to have a negative influence on crime recordings. We like to emphasize that the negative relation of crime with the short dynamics of unemployment (the second factors of the short cycle and irregular components) remain as in our results for the earlier trivariate models. We regard this finding as a a persistent empirical result from our study. Further implications of our study for crime research are discussed in more detail in Section 3.4.

The estimated regression effects and the residual diagnostics for our five-variate model are reported in Table 3.5. We are content with the high levels of goodness of fit that all models have achieved. In particular, Model 3 is relatively parsimonious but the reported R_D^2 are quite satisfactory. Also the reported diagnostic statistics for normality and serial correlation provide sufficient evidence that the model is well specified. The significantly estimated regression coefficients in Model 3 possess the correct signs in all cases. The dummy variables for the Act 1968 and for young males are less convincing while the effects for sentence length, imprisonment and conviction rates clearly lead to lower crime rates.

Table 3.3: Maximum likelihood estimation results for multivariate decomposition model with real GDP, Unemployment and Theft (sample 1955-2001, without and with regression effects)

		Model 1			Model 2			Model 3		
Loadings		Fac1	Fac2		Fac1	Fac2		Fac1	Fac2	
trend	GDP	1	0		1	0		1	0	
	Unemp	−8.1	1		−11.	1		−9.0	1	
	Theft	1.8	.31		2.0	.59		1.9	.39	
cyc-m	GDP	1	0		1	0		1	0	
	Unemp	−4.1	1		−5.4	1		−5.5	1	
	Theft	−2.8	.08		−2.2	.36		−2.3	.29	
cyc-s	GDP	1	0		1	0		1	0	
	Unemp	−14.	1		−9.5	1		−12.	1	
	Theft	1.2	.03		−.03	−.02		.35	−.04	
irreg	GDP	0	0		0	0		0	0	
	Unemp	1	0		1	0		1	0	
	Theft	−.07	1		−.12	1		−.44	1	
Parameters		GDP	Une	Thf	GDP	Une	Thf	GDP	Une	Thf
trend	stdev ×100	.56	5.0	1.2	.51	5.9	1.5	.53	5.2	1.3
cyc-m	stdev ×100	1.2	7.4	3.3	1.0	6.4	2.5	.98	6.5	2.5
	period	10.	10.	10.	12.	12.	12.	11.	11.	11.
	rho	.83	.83	.83	.89	.89	.89	.85	.85	.85
	ampl ×100	1.9	14.	5.8	2.1	20.	7.6	1.7	16.	6.6
cyc-s	stdev ×100	.14	3.1	.19	.18	3.3	.05	.15	3.2	.11
	period	4.5	4.5	4.5	4.6	4.6	4.6	4.7	4.7	4.7
	rho	.96	.96	.96	.96	.96	.96	.95	.95	.95
	ampl ×100	.05	.70	.11	.19	2.2	.44	.14	2.4	.33
irreg	stdev ×100	0	2.3	1.7	0	4.1	1.2	.29	2.9	2.5
Regression effects					GDP	Une	Thf			Thf
Act 68	×100				−1.4	10.	1.9			−3.2
Young M	×100				13.	−51.	−32.			−19.
Sentence	×100				7.3	−1.7	**−50.**			**−34.**
Impris	×100				1.0	−3.3	**−13.**			**−9.7**
Convict	×100				2.8	26.	**−28.**			**−28.**
Diagnostic statistics		GDP	Une	Thf	GDP	Une	Thf	GDP	Une	Thf
N	$\chi^2(2)$	3.4	1.9	1.0	3.4	.60	.31	4.5	1.0	1.8
$Q(13)$	$\chi^2(6)$	11.	7.3	5.7	8.9	6.4	8.3	10.	8.7	9.1
R^2_D		.29	.45	.38	.45	.45	.59	.40	.49	.55

Note: The model includes trend component μ_t in (3.2), medium and short cycles ψ_t in (3.3) and irregular component ε_t as in (3.1). The standard deviations are associated with the disturbances of the different components. For regression effects, values in bold are significant at a 95% significance level. Diagnostics reported are the Bowman-Shanton normality test N, the Ljung-Box test for serial correlation $Q(13)$ and goodness-of-fit criterion R^2_D (compared with fit of random walk plus drift model). The number of observations is 47.

Table 3.4: Maximum likelihood estimation results for multivariate decomposition model with real GDP, Unemployment, Burglary, Theft, and Police (sample 1955-2001, without and with regression effects)

		Model 1		Model 2		Model 3	
Loadings		Fac1	Fac2	Fac1	Fac2	Fac1	Fac2
trend	GDP	1	0	1	0	1	0
	Unemp	−8.8	1	−11	1	−9.7	1
	Burgl	−4.3	.62	−4.0	.84	−3.0	.32
	Theft	−4.6	.45	−.31	.98	.04	.44
	Police	.22	.12	−.03	.18	−.28	.13
cyc-m	GDP	1	0	1	0	1	0
	Unemp	−6.0	1	−6.8	1	−2.6	1
	Burgl	−4.2	.34	−3.8	1.3	−2.7	.21
	Theft	−1.8	−.01	−4.4	3.3	−1.6	.33
	Police	−.30	−.04	.69	−1.1	−.26	−.07
cyc-s	GDP	1	0	1	0	1	0
	Unemp	−9.0	1	−6.6	1	−7.8	1
	Burgl	−1.0	−.16	−3.0	−.03	−3.0	−.06
	Theft	−0.8	−.06	−1.8	−.04	−1.9	−.10
	Police	−.63	.03	−.32	.03	−.49	.03
irreg	GDP	1	0	1	0	1	0
	Unemp	−.52	1	−1.2	1	−2.4	1
	Burgl	−4.4	−.56	−1.4	−.29	−4.0	−.16
	Theft	−2.6	−.40	−1.3	−.12	−2.4	−.17
	Police	.26	.24	−1.0	.16	.05	.14

Note: Table continues on the next page.

Table 3.5: Maximum likelihood estimation results for multivariate decomposition model with real GDP, Unemployment, Burglary, Theft, and Police (sample 1955-2001, without and with regression effects, continued)

		Model 1					Model 2					Model 3				
Parameters		GDP	Une	Bur	Thf	Pol	GDP	Une	Bur	Thf	Pol	GDP	Une	Bur	Thf	Pol
trend	stdev ×100	.54	5.0	3.4	3.2	.28	.64	7.2	3.1	1.5	.28	.66	6.6	2.5	.75	.27
cyc-m	stdev ×100	.83	6.5	3.8	1.5	.31	.14	.96	.54	.69	.14	.54	3.3	1.6	1.3	.26
	period	11	11	11	11	11	7.9	7.9	7.9	7.9	7.9	8.8	8.8	8.8	8.8	8.8
	rho	.91	.91	.91	.91	.91	.98	.98	.98	.98	.98	.92	.92	.92	.92	.92
	ampl ×100	2.1	23	13	3.7	.20	1.2	8.5	4.8	6.1	1.1	1.8	8.6	5.5	4.1	.39
cyc-s	stdev ×100	.22	3.5	.66	.48	.17	.38	4.0	1.2	.84	.18	.35	4.1	1.1	.81	.23
	period	4.6	4.6	4.6	4.6	4.6	4.6	4.6	4.6	4.6	4.6	4.7	4.7	4.7	4.7	4.7
	rho	.96	.96	.96	.96	.96	.95	.95	.95	.95	.95	.94	.94	.94	.94	.94
	ampl ×100	.23	6.5	.95	1.0	.15	.98	2.1	3.4	2.3	.15	.61	2.7	2.0	1.4	.21
irreg	stdev ×100	.45	2.8	3.9	2.1	.77	.54	4.4	2.5	1.6	.87	.50	3.8	3.1	2.0	.74
Regression effects							GDP	Une	Bur	Thf	Pol			Bur	Thf	
Act 68	×100						**−4.3**	20	**39**	4.8	−1.6			**22**	−1.1	
Young M	×100						13	115	31	−44	−1.2			32	−31	
Sentenc B	×100						**−6.8**	**39**	**29**	**18**	−1.5			−2.5	−−	
Impris B	×100						−4.1	**75**	−2.5	−10	4.2			**−12**	−−	
Convict B	×100						−2.8	−28	**−38**	−9.4	−6.1			**−44**	−−	
Sentenc T	×100						**22**	−14	**−66**	**−66**	−7.5			−−	**−25**	
Impris T	×100						**9.1**	**−69**	**−27**	−9.5	−5.8			−−	**−7.6**	
Convict T	×100						**8.3**	33	−28	**−26**	**11**			−−	**−37**	
Diagnostic statistics		GDP	Une	Bur	Thf	Pol	GDP	Une	Bur	Thf	Pol	GDP	Une	Bur	Thf	Pol
N	$\chi^2(2)$	3.9	1.2	.25	.57	4.1	.36	.15	1.9	1.6	8.6	5.1	.79	1.6	.03	7.9
$Q(13)$	$\chi^2(6)$	7.7	7.5	11	8.7	10	6.5	12	11	7.6	12	9.9	14	10	13	3.9
R^2_D		.36	.58	.42	.19	.44	.62	.64	.75	.73	.60	.38	.51	.67	.66	.52

Note: The model includes trend component μ_t in (3.2), medium and short cycles ψ_t in (3.3) and irregular component ε_t as in (3.1). The standard deviations are associated with the disturbances of the different components. For regression effects, values in bold are significant at a 95% significance level. Diagnostics reported are the Bowman-Shanton normality test N, the Ljung-Box test for serial correlation $Q(13)$ and goodness-of-fit criterion R^2_D (compared with fit of random walk plus drift model). The number of observations is 47.

3.4 Literature overview on crime and economic incentives

Many studies into the relationship between economic conditions and crime have been carried out. Some studies rely on unemployment as an indicator of economic conditions (Thornberry and Christenson, 1984; Cantor and Land, 1985; Chiricos, 1987; Cappell and Sykes, 1991; Smith et al., 1992; Britt, 1997; Papps and Winkelmann, 1998; Raphael and Winter–Ember, 2001; Levitt, 2001), some on wages (Grogger, 1998), and a few papers use both unemployment and wages (income) as indicators of economic activity (Ehrlich, 1973; Reilly and Witt, 1996; Entorf and Spengler, 2000; Gould et al., 2002; Machin and Meghir, 2004). An exception is the paper by Ralston (1999) where inflation and unemployment are taken as predictors of crime rates. Literature on the relationship between crime and economic incentives also differs according to the level of aggregation of the employed data: aggregate (national) time series data, aggregate cross-section data, regional panel data, or individual level data. Studies of the first two types often acknowledge the existence of a causal relationship. Further, most literature analyses US and UK data, while for other countries we usually find only a few examples (e.g., Beki et al. (1999) and Boswijk et al. (2002) for the Netherlands; Funk and Kugler (2003a) and Funk and Kugler (2003b) for Switzerland; Raphael and Winter-Ember (2001) for Germany; Fougère et al. (2007) for France; Papps and Winkelmann (1998) and Small and Lewis (1996) for New Zealand). In this section, we review contributions that have shed light on the relationship between crime and economic activity measures. We also relate our empirical findings with the existing empirical literature.

3.4.1 Criminal opportunity and criminal motivation

The seminal paper of Cantor and Land (1985) establishes the macroeconomic link between economic conditions and crime. They propose two paths through which the business cycle could affect the level of crime. *Criminal motivation* works through the impact of changing economic conditions on social strain and social control, whereas *criminal opportunity* affects the availability and vulnerability of criminal targets through routine activities theory. The two effects run in opposition to each other and they need not occur at the same time. For example, a weaker economy will increase the motivation to commit crimes due to greater strain (a pressure individual feels to reach socially determined goals) and reduced social control (the ability of society to regulate its members). Therefore, rising unemployment will have a positive effect on crime through rising criminal motivation for both unemployed (they are out of work and have clear financial motivation) and employed persons (during bad economic times, employed people are often underemployed and they feel pressure to remain in jobs they are unhappy with). On the other hand, a weaker economy also means less opportunities to commit crimes (less jobs, less hours worked, reduced spending and leisure activities, more time spent at home). Cantor and Land (1985) write that an immediate effect of bad economic times and increased unemployment is reduced “circulation of people and property.” The change in criminal opportunity will occur immediately because a weaker economy quickly leads to more unemployed and underemployed people. On the other hand, criminal motivation is a long-run phenomenon and it takes time to develop. For example, those who become unemployed still have resources to rely upon (savings, unemployment benefits), which need to be exhausted before they start feeling financially challenged.[10]

According to Cantor and Land (1985), criminal opportunity means that the level of crime is counter-cyclical in the short-run, falling and rising along with the ups and downs of the cycle in economic activity (Unemployment goes down, Crime goes up; Unemployment goes up, Crime goes down). On the other hand, criminal motivation posits that in the long-run one would expect to find a positive relationship between the level of recorded crime and unemployment (Unemployment goes down, Crime goes down; Unemployment goes up, Crime goes up). It should be noted that Cantor and Land (1985) developed their theory analysing the relationship between crime and unemployment. However, if we follow the opportunity and criminal motivation logic when interpreting the relationship between GDP and crime, we come to an opposite conclusion. Putting these two effects together, one would expect to observe recorded crime behaving pro-cyclically in the short-run, increasing in an immediate response to the positive rates of growth of economies (increase in total volume of goods which present themselves as potential targets for property crime), and behaving counter-cyclically in the long-run, falling in the upswing of an economic cycle and rising on the downswing, as a response to increased strain and decreased social control in the economy. The two effects with respect to GDP and unemployment are summarized in Table 3.6.

Table 3.6: Cantor and Land (1985): Motivational and opportunity crime effects

	Opportunity Effect		**Motivation Effect**	
	Improving economy means more suitable and more attractive targets for criminals in the *short-run*.		Changing economic conditions affect criminal behaviour in the *long-run* through social strain and social control.	
	Bad economy	Good economy	Bad economy	Good economy
GDP	−	+	−	+
Unemployment	+	−	+	−
Crime	−	+	+	−

Note: Crime behaves pro-cyclically in the short-run and counter-cyclically in the long run, when the relationship between GDP and crime is considered. Opposite is true for the relationship between Unemployment and crime.

In order to further clarify the concepts of criminal opportunity and criminal motivation with respect to the relationship between crime and the business cycle, we reference here two more papers. Cook and Zarkin (1985) suggest four potential linkages between the business cycle and crime: (i) legitimate employment opportunities; (ii) criminal opportunities; (iii) consumption of criminogenic commodities (alcohol, drugs, guns); and (iv) response of the criminal justice system. Access to legitimate employment opportunities

[10] For a further discussion on these two effects, we refer interested readers to the original paper by Cantor and Land (1985), and also to Britt (1997) and Paternoster and Bushway (2001).

is pro-cyclical, hence crime recessions would generally promote crime. The quality and quantity of criminal opportunities may be also reduced during recessions (see examples above). Consumption of criminogenic commodities is also pro-cyclical, providing they are normal goods. If these commodities induce criminal behaviour, some crimes would appear pro-cyclical as well. The link between the criminal justice system response to crime and the business cycle is ambiguous. Recession induces less tax revenues which can result in reduced budgets for police and courts and consequently, a reduction in the criminal justice system's capacity to fight crime. On the other hand, recessions might induce potential victims to stay more at home and protect better own and community property, with a consequent increase in probabilities of apprehension in residential neighbourhoods and the capacity of the system to fight crime. The argument is similar to the advantages of lower labour force participations of women with respect to crime (see paper by Witt and Dryden Witte, 1998).

Paternoster and Bushway (2001) give an excellent example of auto theft, motivated by different reasons (joyriding by juveniles and auto theft for profit by 'professionals'), which can behave both pro- and counter-cyclically. Joyriding or theft for temporary use should be immune to financial (motivational) incentives of the business cycle. On the other hand, motivation for theft by professionals should behave counter-cyclically (rising during recession and falling during expansion of the economy). In contract, both type of crimes should be affected by the changes in the criminal opportunity that occur throughout the business cycle. Joyriding may be more attractive during good economic times, because there are more and better cars left unattended, while professional theft may also be more attractive during 'good times' because profits from more expensive cars are higher. In summary, one should be careful when interpreting empirical results one the relationship between crime and the business cycle.

3.4.2 Unemployment and crime

Unemployment is usually a proxy for (lack of) legal income opportunities. Unemployment will make crime more attractive if the alternative is life in poverty. Chiricos (1987) reviews evidence on unemployment-crime relationship, based mainly on the US data, and concludes that evidence favoured the existence of a positive relationship between crime and unemployment. The effect of unemployment is found to be greater on property than on violent crime. Box (1987) also reports 35 studies on the topic, out of which 20 find a positive relationship between crime and unemployment, while the others do not find a significant relationship. More recently, Levitt (2001) summarizes findings of different panel-data approaches, at the level of states, counties, metropolitan areas or cities in the United States, and concludes that "a 1% change in the unemployment rate is typically found to increase property crime by $1-2\%$ contemporaneously but often has no systematic impact on violent crime." Levitt (2001) finds no support for the lagged unemployment rate affecting either property or violent crime, thus suggesting the lack of criminal motivation effect. In what follows, we summarize some important findings that contributed to understanding of the crime-unemployment relationship.

Corman et al. (1987) apply a VAR approach to estimate a relationship between unemployment, arrests, police, demographics and property crimes in New York city from 1970 to 1984. Similarly to this chapter, they correctly assume that the relationship between crime, deterrence (arrests, police) and the business cycle (measured by short-run shocks in unemployment) can have a two-way causation. They conclude that unanticipated jumps in unemployment rates have weak and temporary positive effects on crime, that arrests do deter crimes, and that increases in numbers of police per capita have no impact on either arrests or crime rates. Raphael and Winter-Ember (2001) analyse the relationship between unemployment and crime using the United States state-level panel data, covering the period from 1971 to 1997. Their results consistently indicate that unemployment is an important determinant of property crime rates. Across different model specifications, the authors find that a 1 percentage point decrease in unemployment causes $1-5\%$ decline in property crime. For violent crimes the results are mixed, suggesting a small positive effect of male unemployment rates on state rape rates. The paper by Raphael and Winter-Ember (2001) points out that the available evidence possibly understates the effect of unemployment on crime because of the direction of causation between unemployment and crime. To overcome the possible simultaneity bias, the authors present IV estimation results, using state military contracts and a measure of state exposure to oil shocks as an instrument for unemployment rates. The endogeneity between crime and unemployemnt has also been explored in the paper by Thornberry and Christenson (1984). Thornberry and Christenson (1984) show that unemployment and crime mutually influence one another over the individual's life span, such that unemployment has significant instantaneous effects on crime, while crime primarily has significant lagged effects on unemployment. Using the VAR modelling approach, Saridakis (2004) concludes that unemployment plays a marginal role in violent crime in the US, while income inequality has a significant positive effect on murder. We see this chapter as an extension of literature where crime and economic indicators are treated simultaneously, controlling for the possible two-way causation. In terms of analysed period, our data also cover the 1990s when crime rates in the US and the UK reached record levels, becoming an important public policy issue.

Cappell and Sykes (1991) use US data from 1933 to 1985 to examine the relationship between crime, unemployment, and imprisonment using ARIMA time-series approach. They estimate a simultaneous-equation model, where crime and imprisonment are treated as endogenous variables, and conclude that both the contemporaneous and lagged values of unemployment exhibit modest positive effect on crime. The authors interpret this finding as an evidence for the routine activity and opportunity theory of criminal victimization. Smith et al. (1992) find support both for the criminal motivation and criminal opportunity hypothesis. On the other hand, Britt (1997) only finds support for the criminal motivation effect. Controlling for the variation in the unemployment-crime relationship by age group and over time, he concludes that unemployment has a greater motivational effect on property crime among young adults. A time-varying unemployment-crime relationship is only weekly supported. Neither of the latter two papers include any deterrence or incapacitation controls. Ralston (1999) investigates economic determinants of property crime rates in the United States from 1958 to 1995. He uses inflation and unemployment (structural, cyclical, and frictional) as economic factors that predict property crime rates. Controlling for police and race (black/white) confounders, he concludes that changes in inflation, as well as changes in rates of cyclical and frictional unemployment are important predictors of property crime rates.[11] This chapter extends literature here by proving the existence of both criminal motivation and criminal opportunity effects, while controlling for deterrence and incapacitation confounders.

Summarized literature on the relationship between crime and unemployment is mainly based on the US and UK data. Evidence from other countries are not so abundant. However they also provide mixed results. Entorf and Spengler (2000) use panel data for Germany in the period from 1975 to 1996 to estimate the traditional Becker-Ehrlich deterrence model, controlling for demographic changes, youth unemployment, and income inequality. Using both static and dynamic panel data approaches, which they label 'criminometrics',[12] the authors conclude that "being young and unemployed increases the probability of committing crimes." Fougère et al. (2007) use several sources of data to examine the influence of unemployment on property and violent crime in France in the period 1990 to 2000. They find a positive relationship between unemployment and crime, and that increases in youth unemployment increase crime. They address the potential reverse causality between unemployment and crime by using predicted industrial structure to instrument for unemployment, and conclude that unemployment causally affects burglary, theft, and drug offences. Papps and Winkelmann (1998) use panel data in the period from 1984 to 1996 to examine the relationship between unemployment and different crime categories in New Zealand. Their main finding is that unemployment cannot explain changes in the overall crime rate, although significant effect are found for some categories of crime, like property abuse offences. They find no significant relationship between unemployment and theft, fraud, car conversion, receiving, and burglary, which they label 'dishonesty crimes'.[13] Funk and Kugler (2003b) use data for Switzerland in the period from 1984 to 1998 to estimate a VAR model between theft (robbery), sentence probability, sentence severity, and unemployment. Similarly to Corman et al. (1987), they find that property crimes increase after a shock in unemployment, with a peak response occurring between two and five quarters after the shock.

Evidence from other countries points out the predominance of young people among those committing crimes, suggesting that age is an important factor in explaining crime. With respect to this, possible extension of this chapter would be to control for the variation in the unemployment-crime relationship by age group and over time, in order to see how does the business cycle affect those most likely to engage in crime.[14] Corman et al. (1987) conclude that "more refined data on unemployment by age and race might yield a stronger relationship between crime and the business cycle."

[11]Ralston (1999) writes that "changes in the level of annual cycles of unemployment is positively related to annual changes in the rates of property crimes."

[12]It is interesting to mention here other labels found in the literature for the economics/econometrics of crime. Eide (2000), for example, also uses the term 'criminometric studies'. On the other hand, the Economist of October 2000, gives a generic name of Wackonomics to studies like the economics of economics, the economics of football, and the economics of crime.

[13]In the earlier paper using data on New Zealand and a bivariate estimation approach, Small and Lewis (1996) do find a significant relationship between unemployment and 'dishonesty crimes'.

[14]Crime among young people is often not a consequence of their preferences, but of their poor legal opportunities. As Eide (2000) writes, "young people are perhaps not different, just poorer."

3.4.3 Income and crime

The measures of income represent benefits not only of legal, but also illegal activities. Higher legal income (wages) makes work more attractive than crime, but it also means higher opportunity cost if imprisoned (motivation, negative effect). On the other hand, higher income also means a greater number of more profitable targets for crime (opportunity, positive effect). Eide (2000) lists a variety of income measures found in the crime economics literature: median family income, median income, labour income to manufacturing workers, mean family income, mean income per tax unit, mean income per capita, etc. A large income inequality indicates that crime is a rewarding activity for those at the bottom of the income distribution. In the business cycle context, both income and GDP are indicators of secular economic growth as well as a business cycle. The following literature examines the effect of different income measures on crime.

Ehrlich (1973) uses state-level, cross-sectional data at three points in time and finds that income inequality is positively related to property crimes, while his findings regarding the effect of unemployment are inconclusive.

Cook and Zarkin (1985) examine the effects of the business cycle on robbery, burglary, auto theft and homicide in the US between 1933 and 1981. They use non-parametric and regression approaches in order to isolate the business cycle. In a non-parametric approach, they calculate the average annual rate of growth in crime between each through and subsequent cyclical peak and compare growth rate with the percentage change between the peak year and the subsequent year when the economy is slumping. In a regression approach, they measure effects of the business cycle as deviations from the secular trend in the data, by dividing crime and unemployment or employment rates by the $3-$ or $5-$year moving averages. Both parametric and non-parametric approaches indicate that robbery and burglary behave counter-cyclically, while auto-theft behaves pro-cyclically. Homicide rates are found to be insensitive to business fluctuations. The authors also conclude that effects of secular economic growth may be larger than the effects of cyclical movements. This chapter extends the literature here by explicitly measuring the business cycle using a stationary ARMA process. Similarly to Cook and Zarkin (1985) concluding that effects of secular economic growth may be larger than the effects of cyclical movements, we also find that the shorter cycle of 5 years is less important than the longer cycle of 10 years.

Gould et al. (2002) examine the impact of wages and unemployment on crime using the US data in the period from 1979 to 1997. The authors conclude that both wages and unemployment are significantly related to crime, but that wages played a larger role in the crime trends over the analysed period. They find that the long-term wage trends explain more than 50% of the increase in both property and violent crimes from 1979 to 1997. They also find that increases in unemployment were responsible for much smaller changes in crime rates and that the long-term crime trend has not been affected by the unemployment rate.[15] From the Cantor and Land (1985) perspective, the long-term effect of wages on crime identified by Gould et al. (2002) coincides with the criminal motivation, while the short-term effect of unemployment on crime coincides with the

[15]Compare this result to the finding of Raphael and Winter-Ember (2001) who calibrate that more than 40% of the decline in property crime rates in the US during the 1990s can be explained by the decline in unemployment rates over the same period.

criminal opportunity. The authors correctly assume that economic conditions might be endogenously related to crime and also use instrumental variables (IV) approach, but conclude that "endogeneity is not responsible for the significant relationship between the labour market conditions of unskilled workers and the various crime rates."

Machin and Meghir (2004) examine the effect of changes in wages at the bottom end of the wage distribution on crime rates in England and Wales in the period 1975 to 1996. Their wage measure is the 25^{th} percentile of the real hourly wage distribution. They conclude that worsening the labour market position of less skilled workers increases property crimes like burglary, theft, and vehicle crimes. They also conclude that the labour market opportunities available to individuals are summarized well by the wage variable, therefore unemployment rate plays no role. "Crimes are committed by both the employed and the unemployed. An increase in criminal activity may crowd out leisure as well as work."

Reilly and Witt (1996) use aggregate data from 42 police-force areas in England and Wales over the period 1980 to 1991 to test predictions of Becker's economic model of crime. They find that the growth in male unemployment has a strong positive impact on burglary and theft over the time period considered. The authors estimate that one percent rise in unemployment raises burglary by 0.17% and theft by 0.12%. When it comes to the effect of income, the authors find that per capita household income has a negative effect on the recorded rates of burglary and theft, in the absence of un unemployment variable. The inclusion of the unemployment variable rendered the estimated income effect insignificant. Opposite from Gould et al. (2002) and Machin and Meghir (2004), the authors conclude that the income variable has a proxy role for unemployment and that the effect of economic conditions on crime are fully accounted for by an unemployment measure.

Grogger (1998) examines the relationship between wage rates and criminal participation. He estimates a time-allocation model in which individuals decide how much crime to commit and how much to work as a function of their returns to crime and their wages. Linking crime to wages, Grogger (1998) concludes that youth behaviour is responsive to price incentives and that falling real wages may have been an important factor in rising youth crime in the United States during the 1970s and 1980s. He also emphasizes that wage differentials account for the racial (black/white) differential in criminal participation and the age distribution of crime.

Although summarized literature on income and crime mostly look at this relationship from the perspective of individuals, these findings are also relevant for our macroeconometric approach. As already mentioned, in the business cycle context both income and GDP are indicators of secular economic growth as well as a business cycle. Summarized literature suggests possible further extensions for this chapter, like looking at the effect of a business cycle on crime participation between men and women, across different race categories, and taking into account age distribution of crime. Similarly to the paper by Machin and Meghir (2004), we could also look at the effect between business cycle and income of those at the bottom end of the income distribution, most likely to engage in crimes. An extension here would be to apply macroeconometric approach, advocated in this chapter.

3.5 Conclusion

In this chapter we have analysed macroeconomic time series of real GDP and unemployment simultaneously with burglary rates for England and Wales between 1955 and 2001. By adopting univariate and multivariate structural time series approaches, we have found strong evidence for cyclical behaviour in burglary rates. Further, it is found that the cyclical dynamics in burglary rates can be fully accounted for by cycles of economic activity. We therefore conclude that we have found no evidence of a separate crime cycle, it can be fully explained by economic cyclical behaviour. The presented analysis has accounted for various other variables that can explain variations in burglary. The explanatory variables affect more the short-term dynamics than the long-term dynamics. Further robustness checks in the analysis has been carried out and they do not alter the main conclusion of this chapter. Another interesting finding of this chapter is that the model has detected the motivational and opportunity effects of burglary in relation to unemployment. Although crime theory has discussed these effects, in empirical studies it has been hard to identify such effects. The methodology used in this chapter has been instrumental in the ability to empirically separate these effects of unemployment on burglary and other crimes.

3.A Appendix

3.A.1 Definitions of variables and sources of data

Real GDP : Real gross domestic production per capita in United Kingdom. OECD source, `http://www.sourceoecd.org/`.

Burglary: Number of recorded offences of Burglary (Categories 28 to 31) per capita in England and Wales. Criminal Statistics.

Unemployment: Number registered as unemployed in the UK excluding adult students per capita. Economic Trends.

Conviction Rate (Convict): Number of convictions for burglary in England and Wales divided by the number of recorded burglaries. Criminal Statistics.

Sentence Length (Sentence): Average length (months) of prison sentence for burglary convictions. Criminal Statistics and unpublished data provided by the Home Office.

Prison (Impris): Number imprisoned for burglary divided by number convicted for burglary. Criminal Statistics.

Police (Police force): End of Year Strength (excluding special constables). England and Wales. Annual Abstract of Statistics.

Youths (Young M): Number of males aged 15-24 years as a proportion of the population of England and Wales. Population Trends.

Dummy (Act 68): Theft Act (1968) dummy. $D = 1$ for $t = 1950 - 1968$; $D = 0$ for $t > 1968$.

Chapter 4

Crime and Education: New Evidence From Britain

This chapter is based on work done at the LSE in 2004/2005, under the supervision of Prof. Steve Machin.

4.1 Introduction

In the limited amount of empirical work that exists, education appears to have a significant and large influence on individual propensities to commit crime. For example, in Lochner and Moretti (2004) piece on education and crime in the United States they estimate very large social benefits in terms of crime reductions associated with improved high school graduation rates. They estimate that a 1% reduction in male hugh school drop out rates would save $1.4 billion, or about $2100 per additional male high school graduate. Not a lot is known about the empirical connections between schooling and criminal behaviour in other countries. This chapter attempts to fill this void.

Conceptually, there are difficulties in isolating the impact of education on crime and we spend some time on this in what follows. Specifically, it is difficult to guarantee that the direction of causation flows from education to crime (and not the other way round). To address this problem we adopt a quasi-experimental approach relying on variations in education induced by changes in compulsory school leaving age laws over time to validate the direction of causation.

We look at the relationship between crime and education using two British data sources. The first, the Offenders Index Database (OID) covers all convictions in England and Wales, and we match this to Labour Force Survey data on education for age cohorts over time. The second is data on imprisonment from the 2001 Census, where we can look at crime and education in a large cross-section of the British population.

Our results show sizeable effects of education on crime. Moreover empirical estimates from the instrumental variables strategy that we adopt are, when we use an education variable that is best suited to this approach, rather similar to those that are just based on ordinary least squares regressions that may not have a causal interpretation. In our empirical models of property crime convictions, we report that having low education levels, especially possessing no educational qualifications, is significantly associated with higher levels of offending. We corroborate this with cross-sectional findings on imprisonment and lack of educational qualifications from the 2001 Census. The implications of these findings are clear and they show that improving education amongst offenders and potential offenders should be thought of as a key policy tool in the drive to reduce crime.

The rest of the chapter is organised as follows. Section 4.2 gives some theoretical background on the relationship between education and crime. Section 4.3 describes available crime data sources in Britain and their quality, and gives a brief overview of the schooling reform in Britain in 1944. Section 4.4 describes the data we use in the chapter. Section 4.5 discusses our empirical strategy, whilst Section 4.6 presents the estimation results. Conclusions and ideas for further research are given in the last section of the chapter.

4.2 How education can impact on crime

There are number of theoretical reasons why education may have an effect on crime. According to the existing socio-economic literature, there are several potential channels through which education may have an effect on individuals' criminal behaviour. Lochner (2007) and Feinstein (2002) discuss five reasons why schooling might affect crime: income effects, parenting, peer group effects, pleasure, and patience and risk aversion. For most crimes, one would expect that these factors induce a negative effect of schooling on crime. In what follows, we discuss each of these channels in more detail.

For the case of *income*, education increases the returns to legitimate work, raising the opportunity costs of illegal behaviour. Consequently, subsidies that encourage investments in human capital reduce crime indirectly by raising future wage rates (Lochner, 2004). Additionally, punishment for criminal behaviour often entails imprisonment. By raising wage rates, schooling makes any time spent out of the labour market more costly (Lochner and Moretti, 2004). Therefore, those who can earn more are less likely to engage in crime.[1]

The idea that education raises skill levels and wage rates, which then lowers crime, is not a new one. Ehrlich (1975) empirically examines a number of predictions from an intuitive model relating education to crime. Tauchen et al. (1994) examine the relationship between work, education and crime in a cohort of young men born in 1945 and living in Philadelphia between ages 10 and 18. They find that the time spent at work and in school during a year is negatively, significantly related to the probability of being arrested during the same year. The authors conclude that "the estimated effects of employment and schooling on crime may in part indicate the way in which labour market status affects the opportunity cost of crime." Grogger (1998) examines the relationship between wage rates and criminal participation. He estimates a time-allocation model in which individuals decide how much crime to commit and how much to work as a function of their returns to crime and their wages. The author shows that graduating from high school reduces criminal productivity (although the coefficient is not significantly estimated) and that criminals have on average less education than non-criminals. Linking crime to wages, Grogger (1998) concludes that youth behaviour is responsive to price

[1] For evidence that low wage opportunities correlate strongly with crime see Machin and Meghir (2004), Gould et al. (2002), and Grogger (1998).

incentives and that falling real wages may have been an important factor in rising youth crime during the 1970s and 1980s. He also emphasizes that wage differentials account for the racial (black/white) differential in criminal participation[2] and the age distribution of crime.[3] Machin and Meghir (2004) use panel data on the police force areas of England and Wales between 1975 and 1996 to look at cross-area changes in crime and the low wage labour market, reporting that crime goes down in areas where wage growth in the 25th percentile of the area wage distribution is faster. The authors wished to capture the employment/wage opportunities for the low skill labour market, which is likely to be relevant for those committing property crime, because potential criminals are likely to have low education and labour market attachment. They conclude that "improvements in human capital accumulation through the education system or other means that can be showed to be effective for enhancing individual labour market productivity, coupled with a suitable deterrence measures would be important ingredients in reducing crime."

However, education can also increase the earnings from crime and the tools learnt in school may be inappropriately used for criminal activities. In this sense, education may have a positive effect on crime. Levitt and Lochner (2001) find that controlling for a number of factors (family background, region, ethnicity, etc.), males with higher mathematics scores commit fewer offences, but those with higher scores on mechanical information tests had increased offence rates. Lochner (2004) also writes that "while unskilled crimes should be negatively correlated with education, this need not be true for white collar crimes." He estimates that across cohorts, increases in average education are associated with 11 percent increase in white collar arrest rates (although this coefficient is not significant).

In terms of *parenting*, education could affect parenting skills, which would then have implications on criminal and antisocial behaviour of their children (Rutter et al., 1998). For example, poor child-rearing qualities such as erratic or harsh discipline, serious physical or sexual abuse, low supervision, parental rejection and neglect have been shown to be associated with subsequent criminal involvement and behavioural problems in the offspring. Numerous studies have identified associations between harsh parental punishments and negative behavioural outcomes in children (Smith and Farrington, 2004; Gershoff, 2002; Emery and Laumann-Billings, 1998; Straus, 1994; Parke, 1979). Recent study by Lynch et al. (2006), using data on Australian twins, show a causal relationship between harsh physical punishment and negative behavioural and emotional outcomes in children. It has also been established that offending runs in families (Farrington et al., 1996), whether for coexisting environmental factors or genetic transmission of parental characteristics or a combination of both. Case and Katz (1991) report that youths who had family members in jail when they were being raised are much more likely to be involved in criminal activity. Clearly, cultural factors such as parental expectations, intergenerational learning and family ethics are also important when determining the causes of crime. This

[2]The racial differential in crime rates is in part a labour market phenomenon. Blacks typically earn less than whites, and this wage gap explains about one-fourth of the racial difference in criminal participation rates.

[3]If criminal behaviour responds to wages, then the age distribution of crime is also a labour market phenomenon. In the context of a time-allocation model, this seems quite reasonable. Wages represent the opportunity cost of crime and are well-known to rise with age.

evidence raises the possibility of family-based interventions to reduce subsequent crime.

Education may also be important for teenagers in terms of limiting the opportunity for participating in the criminal activity. "The cynical explanation is that whilst youngsters are at school they are being kept off the streets," (Hansen, 2003). Education can also affect selection of *peer groups*, which may have positive impacts on future behaviour. In the criminological literature it has been shown that delinquent individuals tend to have delinquent friends (Reiss, 1988), that peer effects are greater when people are attached to delinquent peers (Agnew, 1991), and that peer influences have most impact during adolescence (Thornberry and Krohn, 1997). In economics, the empirical evidence suggests that peer effects are very strong in criminal decisions. For example, Case and Katz (1991), using the data from the 1989 NBER survey of youths living in low-income Boston neighbourhoods, find that a 10 percent increase in the neighbourhood juvenile crime rate increases the individual probability of becoming a delinquent by 2.3 percent. The authors also find significant peer effects for drug and alcohol use, and the propensity of youths to be out of school and out of work. Using data from a randomized housing-mobility experiment, Ludwig et al. (2001) estimate that relocating families from high- to low-poverty neighbourhoods reduces juvenile arrests for violent offences by 30 to 50 percent of the arrest rate for control groups. The authors confirm findings in the literature that high-poverty areas have on average more violent but less property crime than low-poverty areas. Calvó-Armengol and Zenou (2004) study the role of social networks and social structure in facilitating criminal behaviour. The authors develop a model of crime decision where delinquents have both positive and negative externality from socializing with each other. Delinquents compete with each other in criminal activities, but benefit from delinquent peers by learning proper know-how on the crime business, that is, by acquiring criminal human capital. They find that otherwise identical individuals connected through a network can end up with very different equilibrium outcomes – either employed, or isolated criminals or criminals in networks. They also show that the aggregate crime level increases with the density of network links.

Pleasure from criminal activity is another channel through which education may have an effect on reducing crime, particularly in the case of juvenile crime. Farrington (2001) reports that when asked their own reasons for criminal participation, teenagers talk about enjoyment whereas older men talk about the material returns to the activity. For example, Clarke and Harris (1992) report that motivation for auto theft differs between juveniles and 'professionals'. Professionals see auto theft as theft for profit, while juveniles see it as theft for temporary use, such as 'joyriding'. Schooling may directly affect these psychological rewards from crime itself. Spear (2000) writes that (under)development of brain maturity in adolescents contributes significantly to manifestations such as propensity to use drugs and participate in risky behaviour. Education may affect the decision to engage in crime by having impact on maturity and development of youths (Hirschi and Gottfredson, 1995). Since most crime is committed by young people and teenagers, the pleasure factor is extremely important and one must attempt to address the question of what role education might play in extenuating this aspect.

Education also influences crime through its effect on *patience* and *risk aversion* (Lochner and Moretti, 2004). Future returns from activities are discounted according to one's patience in waiting for them. Thus, individuals with a lot of patience have low discount rates and value future earnings more highly as compared to those with high discount rates. Oreopoulos (2007) summarizes a sample of literature from psychology and neurology and concludes that young people who drop out of school are myopic, focusing more on immediate costs from schooling (stress from taking tests, uninteresting curricula, forgone earnings, etc.), rather than on future gains from an additional year of schooling. This line of literature also suggests that adolescents lack abstract reasoning skills and are predisposed to risky behaviour. Education can increase patience, which reduces the discount rate of future earnings and hence reduces the propensity to commit crimes. In terms of risk aversion, education may increase risk aversion that, in turn, increases the weight given by individuals to the possible punishment, and hence reduces the likelihood of committing crimes.

In summary, although it is quite hard to quantify the effects of schooling on parenting skills or pleasure, which then have repercussions for crime engagement, as long as schooling increases the marginal return to work more than crime and schooling does not decrease patience levels, we would expect crime to be decreasing in the number of years of schooling. It is also clear that, everything else equal, individuals with higher wage rates and lower discount factors will commit less crime.

4.3 Crime data sources and schooling reform in Britain

4.3.1 Crime data sources

The main source of data on offences committed in England and Wales is *Criminal Statistics*, an annual publication of the Home Office. *Criminal Statistics* publishes a range of tables based on "notifiable offences", which are notifiable in that the police records the crime and notifies the Home Office of their occurrence. Most notifiable offences are "indictable," meaning they can be tried by a judge and jury. However, official statistics are affected by various types of bias and thus can only tell a part of the crime story. It is evident that the true extent of crime will be underestimated, as not every offender is caught, for example. Hence, recorded crime is an imperfect measure of the extent of crime, as it does not include crimes that are not brought to the attention of the police. Further, the official statistics on crime are distorted by variations in the reporting behaviour, police practices (e.g., targeting of certain groups), variations in police recording procedures, and variations in the response made by the judiciary (Rutter et al., 1998). Hence, there exists a serious concern about the "dark figure" of real crime. Other source of crime data in England and Wales is the British Crime Survey (BCS). BCS data are obtained by asking members of the general population about their experiences as victims of crime. Comparing the data in categories that can be compared between the two sources shows that the amount of crime actually committed according to the British Crime Survey may be up to four times the number of crimes reported by the police.

In their paper on crime and economic incentives, Machin and Meghir (2004) use crime data from the official statistics. Comparing official measures to the BCS measure of crime, which some people argue to be a better and preferred measure of actual crime, the authors conclude that "as it is a sample survey, it too potentially suffers from measurement error." In this chapter we use official statistics on offending collected from individuals' criminal records. These data contain police paperwork and detailed information even for offences that did not end up in conviction, and information on the situation surrounding the incident as well as the initial charge. We think that this source of crime data reveals the "dark figure" of real crime in England and Wales better than the *Criminal Statistics*. In their book on antisocial behaviour by young people, Rutter et al. (1998) comment on the quality of the Offender Index data and conclude that "a far better feel for the nature of juvenile delinquency can be achieved by reading such records than by scanning the collated figures."

The third source of crime data in England and Wales is the Census Microdata Samples, which is also not a perfect measure of actual crime. One limitation of the Census data is that they do not differentiate among different types of criminal offences, since we only know if people are imprisoned at the time the Census is taken. From the perspective of this chapter, the second limitation is that we cannot do the credible instrumental variables research design since no data exists that permit us to. The reason is that we only have a single cross-section and, once the full set of age cohort dummies is included in the estimating equation, the instrumental variables are invalid. Instead, we perform the logit estimation approach of the effect of schooling on incarceration rates. It should be noted that in Britain there is hardly any data on crime and education at same time at level of individual. Therefore, an individual-level analysis is not possible given the lack of microdata on both crime and education. This is why we do the cohort approach and match information on offenders from the Offender Index database to education data from the Labour Force Survey.

For the completeness of the discussion on crime data sources, the fourth possible measure of crime are the self-reports on offending. The literature contains extensive discussions of the relative merits of official records versus self-reports of offences (see for example Hindelang et al., 1981 or Weiss, 1986). The methodology employed in such reports is usually to give respondents a standard list of specified criminal activities, as interview questions or as questionnaire for self-completion. However, this measure of crime is also open to a range of biases, which tend to underestimate the true number of offences committed. For example, validity bias is important, because respondents will conceal, exaggerate or forget their offences. There is also interviewer bias, sampling bias, etc.[4]

4.3.2 Background on the schooling reform in Britain

The Education Act of 1944 laid the foundation for education in England and Wales as it is today (see Chan et al., 2002). The Act recognised the importance of education for economic advancement and social welfare. In terms of secondary education, its aim was to provide compulsory secondary education for all children so that every child had equal opportunity to obtain a place in a grammar school, regardless of family background. Most Local Education Authorities (LEA) interpreted the 1944 Act to mean the provision of schooling according to ability. The Act introduced a tripartite system into secondary

[4] See for example David Farrington's report for the Home Office, "What has been learned from self-reports about criminal careers and the causes of offending?", `http://www.homeoffice.gov.uk/rds/pdfs/farrington.pdf`

schools, by using an 'objective' examination, the 11-plus exam (exam taken at around age 11), to test pupils' intelligence and abilities in English and arithmetic. This approach identified three groups of children: (i) Academic pupils, who went to the secondary grammar schools. These schools provided the main route to the university.; (ii) Practical pupils, who went to the technical schools which were vocationally based. There were very few technical schools because the cost of running them was high.; and (iii) Remaining pupils, mostly working class, went to the secondary modern school where they received a more basic education.

The Education Act of 1944 also resulted in the two raisings of the minimum school-leaving age from 14 to 15 in 1947, and from 15 to 16 in 1972. The policy makers' motivation for increasing the school-leaving age was to "improve the future efficiency of the labour force, increase physical and mental adaptability, and prevent the mental and physical cramping caused by exposing children to monotonous occupations at an especially impressionable age" (Oreopoulos, 2006). Harmon and Walker (1995) show that the 1947 change was particularly influential in raising participation in post-compulsory education. That is, many of those who would otherwise have left at the old minimum stayed on beyond new minimum. Oreopoulos (2006) also shows that within two years of the 1947 policy change, the portion of 14-year-olds who left school fell from 57 to less than 10 percent.

The education reform in Britain served as a source of exogenous variation in many papers in labour and health economics. Harmon and Walker (1995) and Oreopoulos (2006) focus on the causal impact of education and earnings. Chevalier et al. (2005), Chevalier (2004), and Galindo-Rueda (2003) look at the effect of parental income on education of their children. Oreopoulos (2006), Doyle et al. (2007), and Lindeboom et al. (2008) examine the impact of education on health.

4.4 Data description and matching approaches

4.4.1 Data

Several sources of data are used in this chapter. The crime and offending data come from the Home Office Offenders Index Data (OID) and the data on imprisonment from the Samples of Anonymised Records (SARs) from the 2001 Census.[5] To convert the OID data into offending rates, we use Office for National Statistics (ONS) population data, and to match other data to the OID (by age cohort, gender and year) we draw on data from the UK Labour Force Survey (LFS) and New Earnings Survey (NES).[6]

4.4.2 Approaches

There are two main empirical approaches we adopt, the first looking at age cohorts from OID data matched to the LFS and NES data sources, the second studying the 2001 Census

[5] Specifically we use the Controlled Access Microdata Samples (CAMS) in the 2001 Census.

[6] The LFS is a large-scale household survey which was carried out in 1975, 1977, 1979, 1981 and then annually from 1983 through 1992, after which it became a quarterly survey. The NES is a 1% employer reported annual survey on individual wages, on which we have access to micro-data from 1975 onwards.

cross-section.

Approach 1: Cohort analysis from OID and LFS/NES

The Offenders Index Database holds criminal history data for offenders convicted of standard list offences from 1963 onwards.[7] The data is derived from the Court Appearances system and is updated quarterly. The Index was created purely for research and statistical analysis. Its main purpose is to provide full criminal history data on selected samples of offenders.

The data set we have access to holds anonymous samples (of about 4 weeks) for each year from the 1960's onwards. The selection of offenders is done by analysis of the court appearance data using the date to select relevant offenders. Selection is based on the following criteria: offenders were chosen where they appeared in court during the first week in March, the second week in June, the third week in September and the third week in November.[8]

The following variables are recorded for each offender: Offenders Index (OI) Number, Date of Birth, Gender, Ethnicity, Appearance Date, Court Code, Curfew Orders, Date of Previous Court Appearance, Age at Appearance, Number of Previous Appearances, Number of Subsequent Appearances, Police Force Code, Offence Class/Sub Class Code, Proceedings Type, Plea, Disposal Codes, and Count of Previous/Subsequent Offences. The main offence groups are Violence against the person; Sexual offences; Burglary; Robbery; Theft and handling stolen goods; Fraud and forgery; Criminal damage; Drug offences; Other (excluding motoring offences); Motoring offences. These are the categories used in most published information that breaks results down by offence category.

For the purposes of this chapter, we extracted offences for the four sampled weeks in each year. It is evident that there is no data on education in the OID. We thus aggregated the offences over age cohorts from 16 to 59 so that we can match to education (and other) data from other sources.[9] In order to match the data with the available information from the Labour Force Survey, we limited the time period from 1984 to 2002. Offending rates (per 1000 population) were calculated using the ONS population data by cohort and year.[10] For the estimation results, criminal offences have been broadly categorised as property crimes (burglary plus theft and handling stolen goods) and violent crimes (violence against the person).

[7]Standard list offences are all indictable or triable offences plus a few of the more serious summary offences. Standard list class codes are set out in the Offenders Index codebook (see Offenders Index (1998a), "Codebook" and Offenders Index (1998b), "A User's Guide," Research Development and Statistics Directorate, Home Office.)

[8]The first week in any calendar month is the week where the Monday is the first Monday in that month.

[9]The data set also contains all offences for the offenders sampled, not just those offences committed during the month sampled. Our concern here is with the latter, but in principle since for each offender, data shows court appearances and offences committed regardless of when these took place, it is possible to analyse patterns of re-offending over time. A good example is the interesting paper by Soothill et al. (2000) who study repeat offending by looking at the criminal records of over 6000 males convicted for sex offences in 1973.

[10]The population data were kindly made available by the UK Office for National Statistics (ONS).

A range of explanatory variables were extracted from the LFS data in the period 1984 to 2002. In particular, we focused on age, date of birth (in order to construct school leaving age dummies), gender, age when completed continuous full-time education, ethnicity, whether employed or unemployed, whether in full-time or part-time work, whether living in London. These variables were first aggregated into cell means by age cohort and year and then matched with the offenders index data in order to form a quasi-panel for age cohorts from 16 to 59 in the period 1984 to 2002. This was done overall and then separately for men and women, and for property and violent crimes. We also carried out the same exercise with data on wages from the New Earnings Survey. Table 4.1 gives summary statistics for the merged Offender Index, Labour Force Survey and New Earnings Survey data.

Table 4.1: Summary statistics, Approach 1

Variable	Sample Size	Mean/Proportion (standard errors)	Min	Max
age	836	37.5 (12.706)	16	59
years of education	836	11.812 (0.721)	9.826	13.705
no qualification (proportion)	836	0.272 (0.157)	0.070	0.753
sla1 (min school leaving age = 15)	836	0.493 (0.492)	0	1
sla2 (min school leaving age = 16)	836	0.449 (0.492)	0	1
property crime rate (per 1000 pop)	836	0.791 (0.877)	0.015	4.212
violent crime rate (per 1000 pop)	836	0.158 (0.167)	0.002	0.696
LFS control variables				
employed (proportion)	836	0.630 (0.081)	0.273	0.754
employed in full-time jobs (proportion)	836	0.758 (0.109)	0.166	0.930
male (proportion)	836	0.490 (0.014)	0.452	0.540
white (proportion)	836	0.936 (0.020)	0.877	0.975
living in London (proportion)	836	0.125 (0.020)	0.081	0.188
NES log hourly wage (up to 2001)	792	2.107 (0.283)	1.022	2.447

Note: Sample period: 1984-2002.

Approach 2: Individual-level analysis from the 2001 Census

The Samples of Anonymised Records (SARs) are samples of individual records from the 1991 and 2001 Censuses. They are micro-data files with a separate record for each individual, covering large sample sizes (between 1-5 percent of the population). The key advantage of the Census data is that we are able to identify individuals who are in prison service establishments (see the Communal Establishment Breakdown in Table 4A.1 in the Appendix 4.A). However, only the 2001 Census has good enough data on individual education and so we are constrained to looking at links between imprisonment and education in the 2001 cross-section only.

The 2001 Individual SAR is a 3 percent sample and contains over 1.5 million records. The Controlled Access Microdata Samples (CAMS) are a more detailed version of the Individual and the Household SAR in 2001, and we can use the CAMS to look at the detailed breakdown of the communal establishment variable (see Table 4A.1) so as to identify prisoners.[11] Similar to the Lochner and Moretti (2004) approach, we want to analyse the impact of education on the probability of incarceration using the UK census data. As already noted though, unlike them, we can only look at one Census cross-section. This has implications for the empirical approach we can adopt, as considered in the next section.

4.5 Empirical approach

4.5.1 The statistical model

For Approach 1, the matched cohort data, consider a simple least squares regression of a measure of offending for a particular age cohort i in year t (O_{it}) with average years of education (E_{it}) as an explanatory variable and X_{jit} ($j = 1, 2, \dots, J$) being a set of other control variables:

$$O_{it} = \alpha_0 + \alpha_1 E_{it} + \sum_{j=0}^{J} \delta_j X_{jit} + u_{it}, \tag{4.1}$$

where u_{it} is an error term in the equation.

If unobserved characteristics of cohorts drive crime participation, but also education, then a least squares estimate of α_1 will be biased. This is the key problem with an empirical model like 4.1 since unobserved characteristics affecting schooling decisions are likely to be correlated with unobservables influencing the decision to engage in crime. For example, α_1 could be estimated to be negative, even if schooling has no causal effect on crime. Individuals who have high criminal returns are likely to spend most of their time committing crime rather than work, regardless of their educational background. As long

[11] While SARs data are available for exploration from the website `http://nesstar.ccsr.ac.uk/nesstarlight/index.jsp`, CAMS data are only available at the ONS offices, after permission has been granted for their use. We are extremely grateful to the CAMS team for giving us access to the data.

as education does not increase the returns to crime, these individuals are likely to drop out of further education. As a result, we might observe a negative correlation between education and crime even though there is no causal effect between the two. Therefore, the challenge is to find an appropriate instrument for education.

To credibly identify a causal impact of education on crime, we adopt a quasi-experimental approach relying on variations in education induced by changes in compulsory school leaving age laws over time to validate the direction of causation. This is like Lochner and Moretti (2004) US study which exploits changes in school leaving age laws across US states. We therefore use two raisings of the school leaving age that occurred in Britain in 1947 and 1973 as instrumental variables in our empirical work. The local variation induced by the two instruments is much more likely to have an impact at the bottom of the education distribution and very little impact at the top of the education distribution, since people near the top would have stayed on after the compulsory school leaving age anyway and the change would not affect them.[12] Therefore, the effect that our empirical approach estimates is the average treatment effect among those who alter their treatment status because they react to the instrument. Imbens and Angrist (1994) call this parameter local average treatment effect (LATE).

Identification is achieved by the inclusion of two dummy variables that record the exogenous change in the minimum school-leaving age (SLA) that occurred in England and Wales in two particular years. In particular, the two dummy variables are defined for individuals who entered their 14$^{\text{th}}$ year between 1947 and 1971 and hence faced a minimum SLA of 15 (variable SLA1), and for those entering their 15$^{\text{th}}$ year from 1972 onwards who therefore faced a minimum SLA of 16 (variable SLA2). The minimum SLA of 14 is our omitted category. Hence we use changes over time in the number of years of compulsory education that government imposed as an instrument for years of education. Since we have more than one instrument, and only one variable to instrument, the model is over-identified, permitting us to implement a two-stage least squares (2SLS) approach.

The set of estimating equations now look as follows:

$$O_{it} = \beta_0 + \beta_1 E_{it} + \sum_{j=0}^{J} \varphi_j X_{jit} + \vartheta_{it}, \tag{4.2}$$

$$E_{it} = \delta_0 + \delta_1 \mathrm{SLA1}_{it} + \delta_2 \mathrm{SLA2}_{it} + \sum_{j=0}^{J} \theta_j X_{jit} + \nu_{it}. \tag{4.3}$$

In this framework, it is important whether changes in compulsory schooling laws act as valid instruments. A legitimate instrument for education in equation (4.1) is a variable that: (i) significantly explains part of the variation in education; and (ii) is not correlated with the unobservables that are correlated with both offending and education. Put alternatively, it is a variable that is a determinant of schooling that can legitimately be omitted from equation (4.1).

To answer the first criteria, let us go back to the definition of our instruments. We use changes over time in the number of years of compulsory education that government

[12] Papers by Lindeboom et al. (2008) and Oreopoulos (2006) show that the first reform in Britain in 1947 only affected the schooling decisions of individuals at the lower end of the education distirbution.

imposed as an instrument for years of education. Harmon and Walker (1995) use same instruments to identify the causal impact of education on wages. They show that the 1947 change was particularly influential in raising participation in post-compulsory education. That is, many of those who would otherwise have left at the old minimum stayed on beyond new minimum. Oreopoulos (2006) shows, using only the 1947 change to instrument for education, that within two years of the 1947 policy change, the portion of 14-year-olds who left school fell from 57 to less than 10 percent.[13]

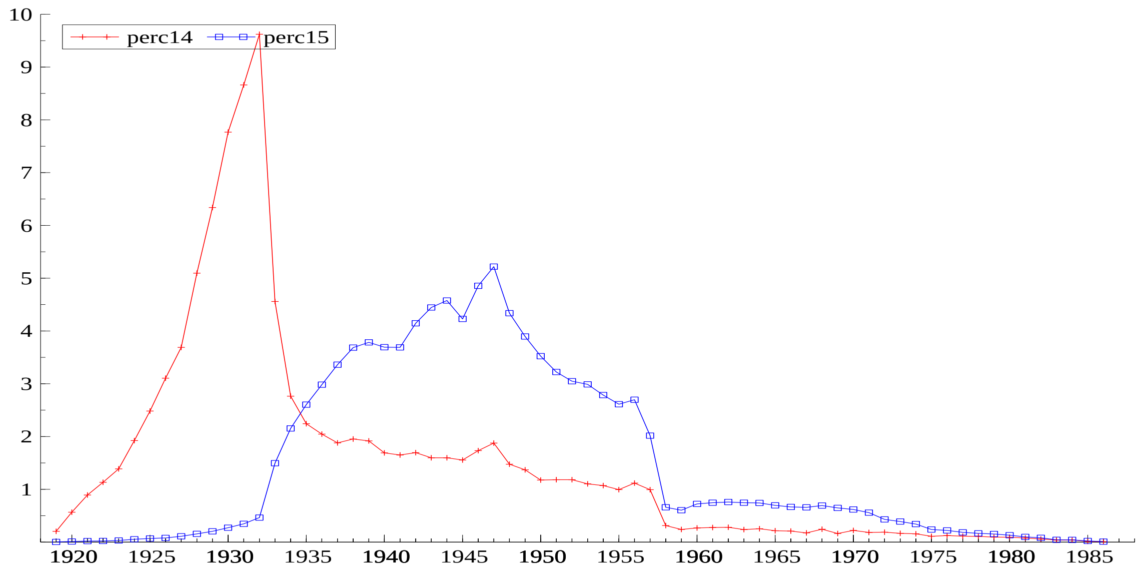

Figure 4.1: Percentage left full-time education aged 14 and 15 by year of birth

Figure 4.1, showing the percentage of those who left full-time education aged 14 and 15 by year of birth, illustrates remarkably the effect of the two legislative changes on education attainment of the affected cohorts. There is a large drop in percentage of those born between 1933 and 1957, who left full-time education aged 14 and who were affected by the first policy change in 1947 (a drop in the cohort of those born in 1932, not affected by the 1947 policy change who left school aged 14, to the cohort of 1933 affected by the 1947 policy change, is 5 percentage points, from 9.6 to 4.6 percentages). There is also a drop in percentage of those born after 1957, who left full-time education aged 15 and who were affected by the second policy change in 1973 (a drop in the cohort of those born in 1957, still not fully affected by the 1973 policy change who mostly left school aged 15, to the cohort of 1958 fully affected by the 1973 policy change, is 1.3 percentage points, from 2.0 to 0.7 percentages). It is interesting to see that there is a further drop in percentage of those who left full-time education aged 14, in the cohorts born after 1957 and affected by the second policy change in 1973. The first stage regressions we report below confirm the strong correlation between the two policy changes and the education measure more formally, both for a years of education variable and for a variable measuring

[13] Oreopoulos (2006) concludes that his IV estimate of the returns to schooling is probably closer to the average treatment effect (ATE) estimator, since the 1947 legislation affected almost half of the population.

lack of educational qualifications.[14]

To answer the second criteria for a valid instrument, we believe our instruments form a plausible identification strategy since changes in compulsory attendance laws have not historically been concerned by problems with crime. To our knowledge, legislators enacting the laws did not act in response to concerns with juvenile delinquency, youth unemployment, or other factors related to crime, thus making schooling laws an appropriate instrument.

4.5.2 Logit models from Census data

For Approach 2, we cannot implement the 2SLS/IV approach in a cross-section like the Census 2001 data since, with a single cross-section, the instruments are simply cohort dummy variables with no cross-time variation to exploit. This renders the instruments invalid if age cohort dummies are included in the estimating equation. Instead, we therefore present logit estimates looking at the association between imprisonment and education at the individual-level, to compare and contrast with the results from our cohort-based approach.

4.6 Estimation results

4.6.1 Approach 1

Figure 4.2 presents the age distribution of property crime convictions per 1000 population, for the selection of years (1984, 1990, 1996, 2002). This "age-offences curve" shows peaks for the age group 16 to 20 and after that a strong tendency to decrease with age. Additionally, peaks seem to be lower in 1996 and 2002 as compared to 1984 and 1990. This is interesting for our empirical analysis since educational attainment rises over time.

Literature abounds with examples of the age-crime profile, that rises until late teens, than falls rapidly (Blumstein et al., 1986; Farrington, 1986; Gottfredson and Hirschi, 1986; Grogger, 1998). This relationship between age and crime is robust over time, across countries, demographic subgroups, types of crime, and holds irrespective of the way the crime is measured (Hirschi and Gottfredson, 1983). The high-to-low pattern of the age-offences curve is also well supported by the underlying theory. The combination of factors mentioned earlier (absence from school, peer pressure, lack of parental guidance, pleasure from committing a crime, etc.), together with the fact that young people tend to be protected from harsh punishment in the criminal justice system, bring about high levels of youth delinquency and criminal activity. As young people grow older, they begin to be influenced by a series of factors which discourage them from breaking the law. As

[14]To formally test for the statistical validity of instruments, we used a test for over-identification, because we have more instruments (SLA1 and SLA2) than endogenous variables (the single education variable on the right hand side of the offences equation). This test failed to reject the null hypothesis that the instruments are valid at the 5% level. We also use the t-test for significance of the instrumental variables (significant at the 1% level) and the F-test for joint significance of the two instruments. Staiger and Stock (1997) use as a rule of thumb that the F-test for joint significance of the instruments in the first-stage regressions should be larger than 10, which our instruments comply with.

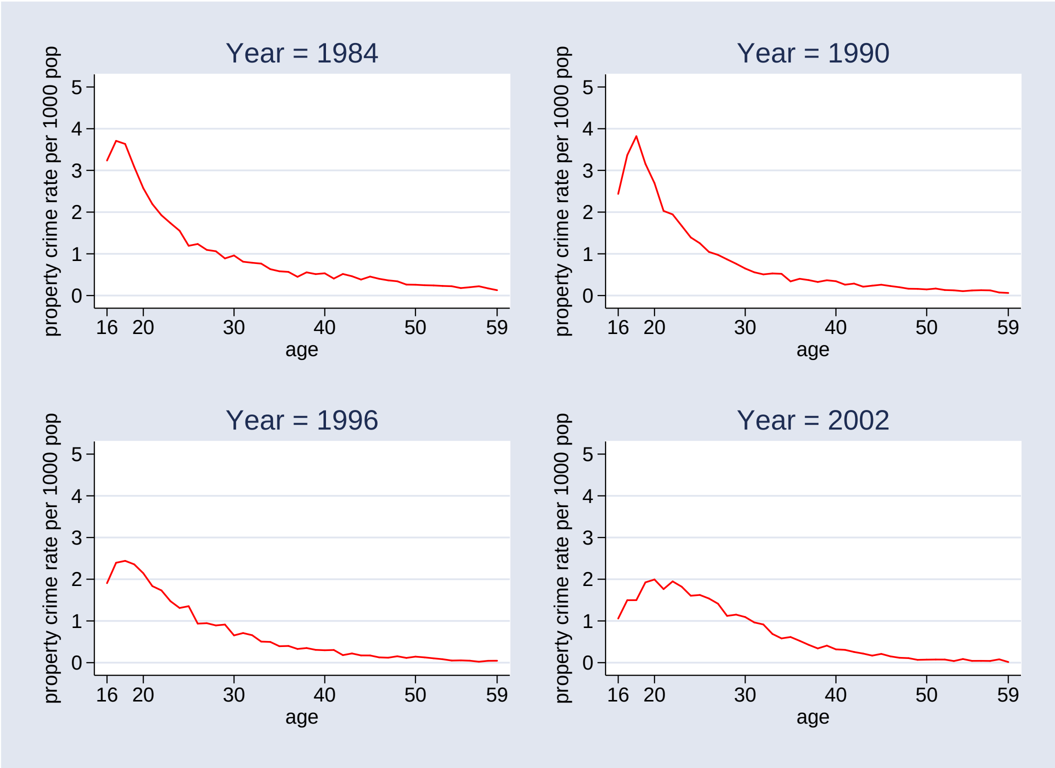

Figure 4.2: Property conviction rates (per 1000 population) by age, 1984 to 2002

they become more independent, enter the labour market, form their own families, and become part of their own communities, young people begin to develop ties to society and attachments to social institutions and norms. These factors, coupled with the possibility of more severe legal sanctions, all encourage a lower crime rate, as young people move towards adulthood (Hansen, 2003). For Grogger (1998), the high-to-low pattern of the age-crime curve is a labour market phenomenon. Wages measure the opportunity cost of crime and grow with age as a worker accumulates labour market experience.

Table 4.2 presents estimation results for different specifications of property crime offending equations, using both the OLS and 2SLS estimating approaches. The upper panel of the Table 4.2 shows six estimates of the link between property crime convictions and average years of education, three estimated by OLS and three by 2SLS. The specifications differ in whether or not they include LFS and NES control variables, so columns (1) and (2) do not, columns (3) and (4) include LFS controls only (detailed in the notes to the Table 4.2) and columns (5) and (6) include LFS controls and the NES hourly wage. The lower panel of the Table shows the first stage years of education equations (for the 2SLS models in columns (2), (4) and (6)).

Considering first the years of education equations in the lower panel of the Table 4.2, it is clear that the first stage is strong, with the instruments being strongly significant (as shown by the F-tests of their marginal significance). For the SLA1 variable the affected cohort has somewhere between 0.47 and 0.50 years more schooling than an unaffected cohort and for the SLA2 variable the comparable numbers are slightly higher at 0.57 to 0.63 years more education. These estimates clearly show that changes in the compulsory

Table 4.2: Property crime convictions and years of education

	Log(Property Crime Convictions Per 1000 Population), by Age and Year, 16-59 Year Olds, 1984-2002					
A. Crime Equations						
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
Years of Education	-0.054	-0.778	-0.153	-0.673	-0.219	-0.606
	(0.055)	(0.124)	(0.054)	(0.106)	(0.050)	(0.093)
Age Dummies (43)	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies (19)	Yes	Yes	Yes	Yes	Yes	Yes
LFS Control Variables	No	No	Yes	Yes	Yes	Yes
NES Hourly Wage	No	No	No	No	Yes	Yes
R^2	0.96	–	0.97	–	0.98	–
Sample Size	836	836	836	836	792	792
B. First Stage Years of Education Equations						
SLA1		0.465		0.498		0.493
		(0.031)		(0.028)		(0.027)
SLA2		0.626		0.568		0.577
		(0.040)		(0.036)		(0.035)
Age Dummies (43)		Yes		Yes		Yes
Year Dummies (19)		Yes		Yes		Yes
LFS Control Variables		No		Yes		Yes
NES Hourly Wage		No		No		Yes
R^2		0.96		0.97		0.98
Sample Size		836		836		792
F-test of Significance of SLA1 and SLA2		$F(2, 772) = 125.6$		$F(2, 767) = 159.7$		$F(2, 723) = 168.0$
(p-value)		($p = 0.000$)		($p = 0.000$)		($p = 0.000$)

Note: Models estimated on age-year cells, for 16-59 year olds between 1984 and 2002. Standard errors in parentheses. The LFS control variables are: proportion male, proportion employed, proportion non-white, proportion of employed in full-time jobs, proportion living in London. SLA1 = 1 for those with compulsory school leaving age of 15 (raised from 14 in 1947), = 0 otherwise; SLA2 = 1 for those with compulsory school leaving age of 16 (raised from 15 in 1973), = 0 otherwise. The specifications including the NES hourly wage are estimated up to 2001 only.

school leaving laws significantly raised the affected cohorts' average years of education.[15]

In the upper panel, the estimated effect of years of education on the rates of offending for property crimes are all estimated to be negative and, with the exception of the column (1) model, are statistically significant. The OLS estimates, in the odd numbered columns, vary from around -0.05 to -0.21, suggesting a 10% increase in the years of schooling lowers property crime convictions by 0.5 to 2%. The 2SLS estimates are bigger (in absolute terms), and lie between -0.60 and -0.77, corresponding to much bigger effects (i.e., a 10% increase in years of schooling lowers crime by 6 to 7.7%). Bearing in mind that a 10% increase in schooling is just over one year, these effects are sizeable. In comparison, estimates from Lochner and Moretti (2004) imply that a one year increase in average schooling levels would reduce both property and violent arrest rates by about 11-12%.

In line with previous studies comparing IV and OLS estimation results, our IV estimates are bigger (in absolute terms) than our OLS estimates. This may seem surprising, considering that OLS estimate of the effect of education is upward biased (in absolute terms), because individuals who have high criminal returns are likely to spend most of their time committing crime rather than work, regardless of their educational background. IV estimate that attempts to correct for this bias should be lower (in absolute terms). One explanation for this is that our measure of education in Table 4.2 assumes linear returns to all levels of schooling. Card (2001) suggests that the higher IV results occur because they approximate average effects among a small group that is probably different from the rest of the sample, whereas OLS estimates approximate average effects among everyone (in the absence of omitted variables and measurement error biases). Oreopoulos (2006) gives a few reasons why LATE estimates from IV could exceed OLS estimates, if we assume that individuals in our sample are rational and they weigh cost and benefits from staying on at school or dropping out and committing crime. It might be that these individuals are more credit constrained, less patient or have greater distaste for school.

In Table 4.3 we consider what is probably a more appropriate education variable, given the nature of the instruments, the proportion of the cohort with no educational qualifications. The resulting OLS estimates are also more similar to the IV results. The table structure is the same as for Table 4.2, with the same six specifications reported in the upper panel, and the first stages given in the lower panel of the Table. Again the first stage works well, with the SLA1 group having a no qualifications proportion around 0.06 to 0.10 lower, and the SLA2 group having a proportion 0.09 to 0.14 lower, depending on specification. Over the sample the mean proportion with no qualifications is about 0.25 so these are sizeable impacts. The F-tests again show the instruments to be strongly significant in the first stage.

We tend to prefer the no qualifications specifications in the 2SLS setting since the 2SLS estimator gives the rate of return for marginal individuals with high discount rates, that is, those with less patience and low education. It is therefore interesting that, in the case of this education variable, the OLS and 2SLS estimates in the upper panel of the Table 4.3 are much more similar to one another. The magnitudes remain large, but also seem more plausible. According to the OLS specifications, a 10 percent increase in

[15]Similar first stage estimation results are to be found in papers by Oreopoulos (2006, 2007), who uses only the 1947 change to instrument for education. He concludes that raising the school-leaving age from 14 to 15 increased the average education by half a year.

Table 4.3: Property crime convictions and no educational qualifications

	Log(Property Crime Convictions Per 1000 Population), by Age and Year, 16-59 Year Olds, 1984-2002					
A. Crime Equations						
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
No Qualifications	3.128	3.215	3.981	2.613	3.340	2.810
	(0.195)	(0.435)	(0.233)	(0.568)	(0.259)	(0.584)
Age Dummies (43)	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies (19)	Yes	Yes	Yes	Yes	Yes	Yes
LFS Control Variables	No	No	Yes	Yes	Yes	Yes
NES Hourly Wage	No	No	No	No	Yes	Yes
R^2	0.97	–	0.98	–	0.98	–
Sample Size	836	836	836	836	792	792
B. First Stage No Qualification Equations						
SLA 1		-0.102		-0.068		-0.062
		(0.008)		(0.006)		(0.005)
SLA 2		-0.141		-0.098		-0.089
		(0.010)		(0.008)		(0.007)
Age Dummies (43)		Yes		Yes		Yes
Year Dummies (19)		Yes		Yes		Yes
LFS Control Variables		No		Yes		Yes
NES Hourly Wage		No		No		Yes
R^2		0.95		0.97		0.98
Sample Size		836		836		792
F-test of Significance of SLA1 and SLA2		$F(2, 772) = 96.8$		$F(2, 767) = 81.5$		$F(2, 723) = 89.1$
(p-value)		($p = 0.000$)		($p = 0.000$)		($p = 0.000$)

Note: Models estimated on age-year cells, for 16-59 year olds between 1984 and 2002. Standard errors in parentheses. The LFS control variables are: proportion male, proportion employed, proportion non-white, proportion of employed in full-time jobs, proportion living in London. SLA1 = 1 for those with compulsory school leaving age of 15 (raised from 14 in 1947), = 0 otherwise; SLA2 = 1 for those with compulsory school leaving age of 16 (raised from 15 in 1973), = 0 otherwise. The specifications including the NES hourly wage are estimated up to 2001 only.

no qualifications corresponds to a 31 to 40 percent higher conviction rate. In the 2SLS specifications, comparable numbers are between 28 percent in the fully specified model (column (6)) and 32 percent for the model with only age and year dummies (column (1)).

It is interesting that this variable is probably more appropriate than the years of education variable in the IV context. As we have already mentioned, this is since the local variation induced by the instrument is much more likely to have an impact at the bottom of the education distribution and no impact at the top since people near the top would have stayed on after the compulsory school leaving age anyway and so the change would not affect them. For these results, the OLS and IV estimates are much more in line with one another and suggest a strong causal link between lack of education qualifications and property crime offending.[16]

Table 4.4 examines the effect of education on convictions for property crimes for the cohort panel of men only. The OLS and 2SLS show the latter to be larger (in absolute terms) for both education variables, although they are in the same ballpark in terms of magnitudes. In the OLS model (column (1)) an additional year of schooling decreases property crime conviction rates by about 26 percent, whereas, for the 2SLS model (column (2)), this number is rather large at 45 percent. On the other hand, a 10 percent increase in no qualifications means a 26 percent higher (OLS model, column (3)) or 35 percent higher (2SLS model, column (4)) conviction rates.

Table 4.5 examines the effect of education on convictions for property crimes for the cohort panel of women only. The OLS and 2SLS show the latter to be larger (in absolute terms) for both education variables, and both OLS and 2SLS results are larger than the estimated effects for men. In the OLS model (column (1)) an additional year of schooling decreases property crime conviction rates by about 30 percent, whereas, for the 2SLS model (column (2)), this number is much larger. On the other hand, a 10 percent increase in no qualifications means a 50 percent higher (OLS model, column (3)) or 76 percent higher (2SLS model, column (4)) conviction rates. The magnitude of the estimated effects between men and women show that the effect of education or lack of education on crime is much larger for women. Literature on crime and women shows that women more often than men drop out of school, which holds them back from being able to get better paying jobs that high school and university graduates may receive. Thus, they may be more tempted to commit crimes to make money to support themselves and their children (many women who quit school also get pregnant at an early age). Another explanation can be that women benefit less from education than men, since the economic rewards from an additional year of education are less marked for women (evidence shows that women are systematically paid less than men). Therefore, they are more likely to drop out of school, ending up in the lower socio-economic standing. Literature also shows that women with the poor socio-economic background are disproportionally represented in prison populations.

So far we have considered only property crime convictions. In Table 4.6 we look at convictions for violent crimes, again amongst males. These results are much less clear regarding the impact of education on crime. In fact violent crimes, if anything, are positively related to education (this is true in the 2SLS models, but effects are estimated

[16] Some economists also argue that the educational degree or the lack of one matters more than the years of schooling. For a discussion, see Card (1999).

Table 4.4: Property crime convictions and no educational qualifications, males separately

	Log(Property Crime Convictions Per 1000 Population), by Age and Year, 16-59 Year Olds, 1984-2001, Males			
A. Crime Equations				
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Years of Education	-0.256	-0.428		
	(0.048)	(0.094)		
No Qualifications			2.549	3.337
			(0.224)	(0.564)
Age Dummies (43)	Yes	Yes	Yes	Yes
Year Dummies (19)	Yes	Yes	Yes	Yes
LFS Control Variables	Yes	Yes	Yes	Yes
NES Hourly Wage	Yes	Yes	Yes	Yes
R^2	0.98	–	0.98	–
Sample Size	792	792	792	792
B. First Stage Equations				
SLA1		0.524		-0.080
		(0.032)		(0.007)
SLA2		0.570		-0.076
		(0.041)		(0.009)
Age Dummies (43)		Yes		Yes
Year Dummies (19)		Yes		Yes
LFS Control Variables		Yes		Yes
NES Hourly Wage		Yes		Yes
R^2		0.97		0.96
Sample Size		792		792
F-test of Significance of SLA1 and SLA2 (p-value)		$F(2, 724) = 133.0$ ($p = 0.000$)		$F(2, 724) = 69.5$ ($p = 0.000$)

Note: Models estimated on age-year cells, for 16-59 year olds between 1984 and 2002. Standard errors in parentheses. The LFS control variables are: proportion male, proportion employed, proportion non-white, proportion of employed in full-time jobs, proportion living in London. SLA1 = 1 for those with compulsory school leaving age of 15 (raised from 14 in 1947), = 0 otherwise; SLA2 = 1 for those with compulsory school leaving age of 16 (raised from 15 in 1973), = 0 otherwise.

Table 4.5: Property crime convictions and no educational qualifications, females separately

	Log(Property Crime Convictions Per 1000 Population), by Age and Year, 16-59 Year Olds, 1984-2001, Females			
A. Crime Equations				
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Years of Education	-0.300	-2.300		
	(0.229)	(0.458)		
No Qualifications			5.035	7.627
			(1.214)	(2.850)
Age Dummies (43)	Yes	Yes	Yes	Yes
Year Dummies (19)	Yes	Yes	Yes	Yes
LFS Control Variables	Yes	Yes	Yes	Yes
NES Hourly Wage	Yes	Yes	Yes	Yes
R^2	0.65	–	0.66	–
Sample Size	792	792	792	792
B. First Stage Equations				
SLA1		0.451		-0.050
		(0.028)		(0.006)
SLA2		0.575		-0.092
		(0.037)		(0.007)
Age Dummies (43)		Yes		Yes
Year Dummies (19)		Yes		Yes
LFS Control Variables		Yes		Yes
NES Hourly Wage		Yes		Yes
R^2		0.97		0.98
Sample Size		792		792
F-test of Significance of SLA1 and SLA2 (p-value)		$F(2, 724) = 138.5$ ($p = 0.000$)		$F(2, 724) = 80.9$ ($p = 0.000$)

Note: Models estimated on age-year cells, for 16-59 year olds between 1984 and 2002. Standard errors in parentheses. The LFS control variables are: proportion male, proportion employed, proportion non-white, proportion of employed in full-time jobs, proportion living in London. SLA1 = 1 for those with compulsory school leaving age of 15 (raised from 14 in 1947), = 0 otherwise; SLA2 = 1 for those with compulsory school leaving age of 16 (raised from 15 in 1973), = 0 otherwise.

to be insignificantly different from zero in the OLS models). As violent crimes do not always respond to economic incentives in the same way as property crimes, these results would need to be interpreted from a different perspective.

However, it is possible that education increases the earnings from crime and the tools learnt in school may be inappropriately used for criminal activities. In this sense, education may have a positive effect on crime. Levitt and Lochner (2001) find that controlling for a number of factors (family background, region, ethnicity, etc.), males with higher mathematics scores commit fewer offences, but those with higher scores on mechanical information tests had increased offence rates. Studies of participants in terrorism, which can be viewed as 'a violent form of political engagement,' often report that majority of individuals involved in terrorist activities are well educated, with some university training (see for example Benmelech and Berrebi, 2007; Krueger and Malečková, 2003; Russell and Miller, 1983). This argument is supported from both the supply and the demand sides. More educated people are more likely to participate in politics, and terrorist organisations may prefer better educated, committed individuals. Benmelech and Berrebi (2007) show evidence that older and more educated suicide bombers are more effective when assigned to important targets, and are more likely to launch successful attacks without being caught. Similarly, Tauchen et al. (1994) find that estimated coefficient on IQ is negatively related to the probability of being arrested. They conclude that "this is consistent with the hypothesis that individuals with higher IQ-s have better legal labour market opportunities, but may also reflect the ability of high intelligence individuals to avoid arrest." It is possible to argue here that violent crimes are complex tasks which require a considerable level of task specific and general human capital. If policy makers attempt to use education as part of the strategy to reduce violent crime, they should not limit themselves to increasing years of schooling, but also consider the content of education, insisting on good instructions in mathematics and science (see for example Krueger and Malečková, 2003).

The final set of cohort results we consider are some robustness checks, given in 4.7. Here we present results for two specific age sub-sets, again for men. First, in columns (1) and (2) we look at cohorts of men aged 21-59, dropping younger cohorts who have not completed their education. Of course, this matters much more for the years of schooling measure and so it is perhaps not surprising that the estimated effects for the no qualifications variable are not much affected. Finally, we look at the age 21-40 cohorts (in columns (3) and (4)), again finding the main results to be highly robust. From this we conclude that our OID cohort analysis uncovers an important causal link between possessing no educational qualifications and the probability of committing property crime offences.

4.6.2 Approach 2

Evidence from the Offenders Index data is consistent with the hypothesis that higher education levels reduce convictions for engaging in criminal activity. We corroborate these findings using Census 2001 data. One limitation of the Census data is that they do not differentiate among different types of criminal offences, since we only know if people are imprisoned at the time the Census is taken. The second limitation is that we cannot implement the 2SLS approach since we only have a single cross-section and, once the

Table 4.6: Violent crime convictions and no educational qualifications, males separately

	Log(Violent Crime Convictions Per 1000 Population), by Age and Year, 16-59 Year Olds, 1984-2001, Males			
A. Crime Equations				
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Years of Education	0.043	0.335		
	(0.067)	(0.130)		
No Qualifications			-0.536	-1.402
			(0.331)	(0.828)
Age Dummies (43)	Yes	Yes	Yes	Yes
Year Dummies (19)	Yes	Yes	Yes	Yes
LFS Control Variables	Yes	Yes	Yes	Yes
NES Hourly Wage	Yes	Yes	Yes	Yes
R^2	0.96	–	0.96	–
Sample Size	792	792	792	792
B. First Stage Equations				
SLA1		0.524		-0.080
		(0.032)		(0.007)
SLA2		0.570		-0.076
		(0.041)		(0.009)
Age Dummies (43)		Yes		Yes
Year Dummies (19)		Yes		Yes
LFS Control Variables		Yes		Yes
NES Hourly Wage		Yes		Yes
R^2		0.97		0.96
Sample Size		792		792
F-test of Significance of SLA1 and SLA2 (p-value)		$F(2, 724) = 133.0$ ($p = 0.000$)		$F(2, 724) = 69.5$ ($p = 0.000$)

Note: Models estimated on age-year cells, for 16-59 year olds between 1984 and 2002. Standard errors in parentheses. The LFS control variables are: proportion male, proportion employed, proportion non-white, proportion of employed in full-time jobs, proportion living in London. SLA1 = 1 for those with compulsory school leaving age of 15 (raised from 14 in 1947), = 0 otherwise; SLA2 = 1 for those with compulsory school leaving age of 16 (raised from 15 in 1973), = 0 otherwise.

Table 4.7: Checks of robustness

	Log(Property Crime Convictions Per 1000 Population), by Age and Year, 1984-2001, Males			
	Ages 21-59		Ages 21-40	
A. Crime Equations				
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
No Qualifications	2.610	2.850	2.260	3.448
	(0.271)	(0.796)	(0.290)	(0.792)
Age Dummies (38/19)	Yes (38)	Yes (38)	Yes (19)	Yes (19)
Year Dummies (19)	Yes	Yes	Yes	Yes
LFS Control Variables	Yes	Yes	Yes	Yes
NES Hourly Wage	Yes	Yes	Yes	Yes
R^2	0.97	–	0.98	–
Sample Size	702	702	360	360
B. First Stage No Qualification Equations				
SLA1		-0.064		–
		(0.007)		–
SLA2		-0.061		-0.031
		(0.009)		(0.004)
Age Dummies (38/19)		Yes (38)		Yes (19)
Year Dummies (19)		Yes		Yes
LFS Control Variables		Yes		Yes
NES Hourly Wage		Yes		Yes
R^2		0.97		0.96
Sample Size		702		360
F-test of Significance of SLA1 and SLA2 (p-value)		$F(2, 639) = 41.8$ ($p = 0.000$)		$F(1, 317) = 52.0$ ($p = 0.000$)

Note: Models estimated on age-year cells, for 16-59 year olds between 1984 and 2002. Standard errors in parentheses. The LFS control variables are: proportion male, proportion employed, proportion non-white, proportion of employed in full-time jobs, proportion living in London. SLA1 = 1 for those with compulsory school leaving age of 15 (raised from 14 in 1947), = 0 otherwise; SLA2 = 1 for those with compulsory school leaving age of 16 (raised from 15 in 1973), = 0 otherwise.

full set of age cohort dummies is included in the estimating equation, the instrumental variables are invalid. Instead, we perform the logit estimation approach of the effect of schooling on incarceration rates.

Table 4.8 presents summary statistics for imprisonment rates, for both men and women, by different age groups. Overall, 0.13 percent of the 16-64 year olds in the British population were in prison, according to the 2001 Census. Imprisonment rates for young men aged 16 to 20 are higher than average at 0.34 percent, and reach their highest at 0.57 percent among the age 21 to 25 males. The imprisonment rates then declines for older age groups. This is in line with the postulations of the "crime-age curve" that we saw earlier, when using the Offenders Index data. Far fewer women are in prison and even amongst the highest sub-group (again aged 21-25) imprisonment rates remain low.

The rest of the Table 4.8 shows stark differences by education. Table 4.8 also presents the gap in imprisonment rates between those with no and some qualification. Amongst some of the no qualification groups, the percent in prison rises sharply. For example, 2.57 percent of men aged 21-25 with no educational qualifications are in prison in 2001. The last two columns of the Table show imprisonment gaps between the no qualification and some qualification groups. The gaps are reported in two ways, as percentage gaps and as relative risks. It is evident that there are large gaps in imprisonment rates that are related to the possession of educational qualifications. Moreover, the gaps are at their largest for the age groups where more people are in prison: see the largest relative risk ratios in the final column for the age 21-25 group, for both men and women.

Table 4.9 presents logit estimates that condition upon an additional range of variables from the Census (listed in the Table). Two sets of specifications are reported for the whole sample size, men and women separately, and for the different age by gender groups. The two specifications differ in the way in which they specify the age control variable - in column (1) it is entered as a quadratic, in column (2) a full set of age dummies is included. The logit regression model is based on the log odds ratio ($\log[p/(1-p)]$, where p is a 0/1 variable indicating whether a person is in prison or not), which represents the probability of a success compared with the probability of failure. Hence, an interpretation of estimated coefficients in the logit regression which is usually more intuitive is the 'odds ratio' or the relative risk ratio (RRR), reported in the Table 4.9, together with standard errors and marginal effects. The marginal effects, reported in the square brackets, show the change in the probability of imprisonment due to a unit change of an education variable in question. Logistic regression does not have an equivalent to the R^2 that is found in OLS regression. The goodness-of-fit is based upon likelihood ratio test and McFadden's pseudo R^2. Usually, goodness-of-fit is fairly low for discrete choice models. For a general discussion of alternative goodness-of-fit measures in non-linear models, see Cameron and Trivedi (2005), Section 8.7.

The results in Table 4.9 confirm the descriptive analysis. Even after conditioning on a range of factors, there is a sizeable gap in imprisonment rates between those with no qualifications and those with some educational qualifications. For the full sample, the RRR of around 4 shows that people with no qualifications are four times more likely to be in prison than those with some qualifications. For young men these odds rise even

Table 4.8: Imprisonment rates (percent), 2001 Census

	All		No Qualifications		Some Qualifications		Gap	RRR
	IR	# People	IR	# People	IR	# People	(Std. Err.)	
All	0.13	1099639	0.23	294871	0.09	804768	0.14 (0.01)	2.58
Men	0.25	535820	0.44	142373	0.17	393447	0.27 (0.02)	2.61
Men, Aged 16-20	0.34	57418	0.91	12048	0.18	45370	0.73 (0.06)	5.03
Men, Aged 21-25	0.57	52889	2.57	6176	0.30	46713	2.27 (0.10)	8.67
Men, Aged 26-30	0.42	58977	1.41	8036	0.27	50941	1.14 (0.08)	5.37
Men, Aged 31-64	0.16	366536	0.22	116113	0.13	250423	0.09 (0.01)	1.74
Women	0.01	563819	0.03	152498	0.01	411321	0.02 (0.003)	2.70
Women, Aged 16-20	0.01	54934	0.05	9856	0.01	45078	0.04 (0.01)	7.62
Women, Aged 21-25	0.04	53852	0.17	5229	0.02	48623	0.15 (0.03)	8.38
Women, Aged 26-30	0.02	62415	0.07	7153	0.01	55262	0.06 (0.02)	4.83
Women, Aged 31-64	0.01	392618	0.02	130260	0.01	262358	0.01 (0.003)	2.22

Note: Based on 16-64 year olds in the 3% Census microdata sample. IR = Imprisonment rate; Gap = Gap in imprisonment rate between no and some qualifications; RRR = relative risk ratio.

Table 4.9: Logit estimates of imprisonment equations

	Specification 1		Specification 2	
	No Qualifications	Sample Size Goodness-of-fit	No Qualifications	Sample Size Goodness-of-fit
All	1.389 (0.058)	1099639	1.417 (0.058)	1099639
	[0.18]	$LL = -8914.55$	[0.18]	$LL = -8874.29$
	RRR = 4.00	$R_p^2 = 0.17$	RRR = 4.12	$R_p^2 = 0.17$
Men	1.385 (0.060)	535820	1.412 (0.060)	535820
	[0.34]	$LL = -8206.22$	[0.35]	$LL = -8164.81$
	RRR = 4.00	$R_p^2 = 0.11$	RRR = 4.11	$R_p^2 = 0.11$
Men, Aged 16-20	2.209 (0.152)	57418	2.210 (0.152)	57418
	[0.74]	$LL = -1142.62$	[0.74]	$LL = -1142.58$
	RRR = 9.11	$R_p^2 = 0.12$	RRR = 9.11	$R_p^2 = 0.12$
Men, Aged 21-25	2.011 (0.122)	52889	2.011 (0.122)	52889
	[1.14]	$LL = -1657.21$	[1.14]	$LL = -1657.21$
	RRR = 7.47	$R_p^2 = 0.11$	RRR = 7.47	$R_p^2 = 0.11$
Men, Aged 26-30	1.302 (0.143)	58977	1.301 (0.144)	58977
	[0.54]	$LL = -1439.08$	[0.54]	$LL = -1438.80$
	RRR = 3.67	$R_p^2 = 0.10$	RRR = 3.67	$R_p^2 = 0.10$
Men, Aged 31-64	0.716 (0.092)	366536	0.717 (0.092)	366536
	[0.11]	$LL = -3828.89$	[0.11]	$LL = -3810.81$
	RRR = 2.05	$R_p^2 = 0.10$	RRR = 2.05	$R_p^2 = 0.11$
Women	1.464 (0.254)	563819	1.498 (0.254)	563819
	[0.02]	$LL = -692.74$	[0.02]	$LL = -667.85$
	RRR = 4.32	$R_p^2 = 0.12$	RRR = 4.47	$R_p^2 = 0.15$
Women, Aged 16-20	2.694 (0.755)	54934	2.697 (0.754)	54934
	[0.04]	$LL = -68.55$	[0.04]	$LL = -67.56$
	RRR = 14.79	$R_p^2 = 0.13$	RRR = 14.84	$R_p^2 = 0.13$
Women, Aged 21-25	2.095 (0.510)	53852	2.097 (0.510)	53852
	[0.07]	$LL = -148.51$	[0.07]	-148.45
	RRR = 8.12	$R_p^2 = 0.13$	RRR = 8.14	$R_p^2 = 0.13$
Women, Aged 26-30	0.886 (0.697)	62415	0.878 (0.699)	62415
	[0.02]	$LL = -100.92$	[0.02]	$LL = -100.92$
	RRR = 2.42	$R_p^2 = 0.18$	RRR = 2.41	$R_p^2 = 0.18$
Women, Aged 31-64	1.150 (0.346)	392618	1.157 (0.346)	392618
	[0.01]	$LL = -355.63$	[0.01]	$LL = -355.63$
	RRR = 3.16	$R_p^2 = 0.13$	RRR = 3.18	$R_p^2 = 0.13$

Note: Specification 1: Includes quadratic in age, 15 country of birth dummies, gender dummy (where applicable), non-white dummy, 5 marital status dummies, dummy for never worked, dummies for country; Specification 2: Includes age dummies, 15 country of birth dummies, gender dummy (where applicable), non-white dummy, 5 marital status dummies, dummy for never worked, dummies for country. Heteroscedastic consistent standard errors in round brackets, marginal effect × 100 in square brackets; RRR = relative risk ratio; LL = log-likelihood; R_p^2 = McFadden's pseudo R^2.

more, to around 9.1 for 16-20 year olds, and to 14.8 for women in the same age group.[17] In comparison, estimates in Lochner and Moretti (2004) suggest that, on average, extra year of schooling reduces the probability of imprisonment by slightly more than 0.1 percentage point for whites and and by about 0.4 percentage points for blacks.

4.7 Conclusions

Education can affect the likelihood of offending in a variety of ways. In many theoretical approaches more education leads causally to lower crime. For example, amongst young people staying at school rather than on the streets may well influence their choice of peers and enforce some level of discipline upon them. More generally, education encourages people to develop skills and acquire knowledge and training that affects their future life chances, like acquiring legally paid jobs with satisfactory wages.

This chapter presents some new evidence on the effect of education on crime, looking at two different data sources from Britain. In the first, property crime convictions in England and Wales (taken from using unique Home Office Offenders Index data matched to household and employer surveys) are seen to be significantly lower amongst age cohorts where education is higher. To ensure the direction of causation runs from education to crime (and not in the opposite direction), we follow the idea of Lochner and Moretti (2004) and use exogenous changes in school leaving age laws as instruments for our education variable. Two stage least squares estimates based on this approach show that education significantly reduces property crime convictions among both men and women. Our second approach, based on using 2001 Census data on imprisonment and educational qualifications, corroborate these finding by demonstrating that having no educational qualifications significantly increases the risk of imprisonment. Based on our estimates, government should develop strategy which would target young people with no or very law educational attainment in order to reduce offending and imprisonment.

As always there is a range of possible extensions of this work. For example, further analysis could usefully break down among different types of crime in more detail, examine factors that affect the crime-age profiles of those with some and no educational qualifications, and try to estimate social savings from crime reduction.

There is little doubt that findings from this chapter have important implications for longer-term efforts aimed at reducing crime. For example, a policy that subsidizes schooling and human capital investment is likely to reduce crime in the long run by increasing skill levels. Hence, improving education amongst offenders and potential offenders should be a key policy tool in the drive to reduce crime. Ralston (1999) concludes that "proba-

[17]The Census education variable is more detailed than the no/some educational qualifications split we consider. There is information on five qualification levels, ranging from Level 0 (No Qualifications) through to Level 4 (Degree or higher). We look at the no/some distinction so we can include the young people in our sample since some may not have completed their education, and these are an important group to consider in studies of criminal activity. Specifications estimated for older samples that enter in four dummy variables for No Qualifications, Level 1, Level 2 and Level 3 (omitting Level 4 as the reference category) show a monotonic relationship between the probability of imprisonment and qualification attainment. For example, for men aged 26-30 the relative risk ratios were estimated as 13.46 (Level 0), 6.32 (Level 1), 5.56 (Level 2), 2.27 (Level 3).

tioners and former inmates must be allowed to establish some viable link to the 'legal' economic infrastructure before they are cast afloat in the environment in which they developed the criminal tendencies in the first place."

It is evident that crime is a negative externality with enormous social costs. If education reduces crime, then schooling will have social benefits that are not taken into account by individuals. In this case, the social returns to education may exceed the private return. Hence, policy makers should continue to devote attention towards the design of social and educational policies that can have an impact on the crime-education relationship.

4.A Appendix

Table 4A.1: Type of communal establishment, England, Wales and Scotland, Census 2001

Value	Label	Percentage
-9	Not applicable	98.3
1	NHS psychiatric hospital	0.0
2	Other NHS hospital/home	0.1
3	LA Children's home	0.0
4	LA Nursing home	0.0
5	LA Residential care home	0.1
6	LA Other home	0.0
7	HA home or hostel	0.0
8	Nursing homes (not HA/LA)	0.3
9	Residential home (not HA/LA)	0.4
10	Children's home (not HA/LA)	0.0
11	Psychiatric hospital (not HA/LA)	0.0
12	Other hospital (not HA/LA)	0.0
13	Other medical and care home (not HA/LA)	0.0
14	Defence establishment (inc. ships)	0.1
15	**Prison service establishment**	**0.1**
16	Probation/bail hostel (not Scotland)	0.0
17	Educational establishment	0.5
18	Hotel/boarding house, guest home	0.1
19	Hostel (inc. youth hostel, hostels for homeless and persons sleeping rough)	0.1
20	Civilian ship, boat or barge	0.0
21	Other	0.1

Note: Source: 2001 Individual CAMS Codebook, http://www.ccsr.ac.uk/sars

Chapter 5

Why Are Criminals Less Educated than Non-Criminals? Evidence from a Cohort of Young Australian Twins

This chapter is based on work done at the CPB during 2006–2008, under the supervision of Dr. Dinand Webbink and Dr. Pierre Koning.

5.1 Introduction

Many studies document a strong negative association between education and crime. For instance, in the US two-thirds of all incarcerated men in 1993 had not graduated from high school (Freeman, 1996). Studies that use self-reported and (administrative) arrest data find large differences in property and violent crime across education groups (Tauchen et al., 1994; Lochner, 2004). However, the relationship between crime and education is not straightforward. Does crime reduce investment in human capital or does education reduce criminal activity?

It is difficult to empirically establish the relationship between education and crime because the decisions to participate in crime or invest in human capital can be affected by a multitude of unobserved factors. A major difficulty in studying the relationship between human capital and crime is to find credible exogenous variation in either human capital or in criminal behaviour. Typically, a person's level of schooling is not randomly determined but the result of individual choices. Highly educated individuals might also have unobserved factors that prevent them from committing crimes. Unobserved factors that are both correlated with the decision to invest in human capital and the decision to participate in crime will confound the empirical relationship between education and crime. A second difficulty in studying the relationship between human capital and crime is reverse causality. It is well known that crime rates typically increase with age during the teenage years and decline afterwards. This early criminal involvement might be detrimental for the accumulation of human capital. In particular, teenagers may spend less time in attending school and may consider the future benefits of a criminal career more lucrative than the opportunity costs of schooling.

This chapter studies the relationship between human capital and crime using data of a sample of young Australian twins. We exploit two aspects of the Australian survey data on education and crime. First, because the data are obtained from twins, we are able to control for many unobserved characteristics affecting both criminal behaviour and schooling decisions. Second, because criminal behaviour is measured over different periods of time – prior to and after senior high school completion – we can address the causality between crime and education as well. Since early criminal behaviour may affect human capital formation, and human capital may influence criminal behaviour in later stages of life, we follow a two step analysis.

First, we study the relationship between early crime and the accumulation of human capital. In particular, we estimate the effect of arrests before the age of 18 on educational attainment by using within-twin estimation. In addition, we investigate whether the timing of the arrest matters for educational attainment. Second, we estimate the effect of educational attainment on three measures of crime: incarceration, arrests since the age of 18, and number of arrests. As early criminal behaviour might be an important confounder in the estimation, we control for early arrests and measures of conduct disorder within pairs of twins.

This chapter contributes to the economic literature on the relationship between education and crime in several aspects. First, the empirical economic literature on human capital and crime that takes unobserved factors into account is limited. Two previous studies use arguably exogenous variation in human capital to investigate the effect of education on crime (Lochner and Moretti, 2004; Machin and Vujić, 2008 (chapter 4 of this dissertation)). Both studies use changes in compulsory schooling laws as an instrument for educational attainment, in an attempt to show that education reduces crime. We add to this literature and use an identification strategy that has not been applied before – that is, we exploit the longitudinal nature of our data in order to estimate the relationship between human capital and crime in both directions. Second, we investigate the effect of early criminal behaviour on investment in human capital while controlling for fixed effects within pairs of twins. We are not aware of studies in the economic literature that estimate the causal effect of early criminal activity on educational attainment. Third, there is growing interest in the economic literature over the effects of early conditions in life on adult outcomes (Currie and Stabile 2006, 2007; Borghans et al., 2008). This chapter addresses similar issues.

We find early arrests (arrests before the age of 18) to have a strong effect on human capital accumulation. In particular, early arrests reduce educational attainment with 0.7 to 0.9 years of education and lower the probability of completing senior high school with 20 to 23 percentage points. These effects are largely driven by the timing of the early arrest, showing that arrests at age 13, 14 or 15 are most detrimental for educational attainment. We also find human capital to reduce crime. Completing senior high school reduces the probability of incarceration with 2 to 3 percentage points. Similar but statistically insignificant effects are obtained for the probability of being arrested since the age of 18 and for the number of arrests. When controlling for early arrests and early behaviour problems, the estimated effect of human capital on crime reduces to less than a quarter of the previously estimated association. The strong detrimental effects of early criminal behaviour become also transparent if we consider the estimated effects of early arrests on all three measures of crime. We then find large effects of early criminal behaviour on participation in crime later on. These effects are much larger than the (isolated) impact of human capital on crime. We conclude that the strong association between human capital and crime is mainly driven by the effect of early criminal behaviour on educational attainment. Programs that succeed in preventing early criminal behaviour might yield high social and private returns.

The remainder of this chapter is organised as follows. The next section reviews previous studies. Empirical strategy is presented in Section 5.3, while Section 5.4 describes the data. Section 5.5 shows estimation results which investigate the effect of early arrests on educational attainment. Section 5.6 shows estimation results which analyse the effect of human capital on crime. Robustness checks are presented in Section 5.7, while Section 5.8 decomposes the association between crime and education. Section 5.9 concludes and discusses some policy implications.

5.2 Previous studies

The major difficulty in studying the relationship between human capital and crime is that both variables are driven by a multitude of unobserved factors. For instance, a person's level of schooling is typically not randomly determined but the result of ability and individual choices. These individuals might also have unobserved factors that prevent them from committing crimes. Unobserved factors that are both correlated with the decision to invest in human capital and the decision to participate in crime will confound the empirical relationship between education and crime. As such, OLS estimates of the effects of human capital on crime or OLS estimates of the effects of crime on human capital are likely to be biased.

The first part of this chapter focuses on the effect of early criminal behaviour on human capital formation. To our knowledge there are no previous economic studies that empirically estimate the effect of early crime on investment in human capital while taking unobserved factors into account. Related studies can be found in health economics. Some recent studies investigate the effect of childhood mental health problems such as attention deficit hyperactivity disorder (ADHD), aggression, anti-social behaviour, and depression on human capital accumulation later in life (Le et al., 2005; Currie and Stabile 2006, 2007; Fletcher and Wolfe, 2008; Slade and Wissow, 2007). These studies typically find large negative effects of childhood mental health problems on educational attainment. Other source of related literature focuses on the importance of cognitive and non-cognitive skills for labour market outcomes and social behaviour (Borghans et al., 2008; Carneiro and Heckman, 2003; Heckman et al., 2006; Blanden et al., 2006; Heckman and Masterov, 2007). These studies stress the importance of skill development early in life for human capital accumulation and social success at the later stages in life. Early schooling programmes, like the Perry Preschool Programme (PPP), the Syracuse Programme (SP) or the Head Start Programme (HSP) have proven to be highly effective in reducing criminal activity, promoting socioeconomic skills, and integrating disadvantaged children into mainstream society (see for instance Schweinhart et al., 1993; Donohue and Siegelman, 1998; Lally et al., 1988; Garces et al., 2002). These social, motivational, and emotional skills affect performance in school and in the workplace. Programmes that aim at intervening in the lives of children in their teenage years only attempt to redress the damage of bad childhoods (Carneiro and Heckman, 2003).

The second part of this chapter studies the causal effect of human capital on crime. So far, only two papers in the economic literature try to establish a causal relationship between education and crime (Lochner and Moretti, 2004; Machin and Vujić, 2008 (chapter 4 of this dissertation)). Both studies use changes in compulsory school leaving age laws in order to account for the endogeneity of schooling decisions. Using US Census data, Lochner and Moretti (2004) show that one more year of schooling reduces the probability of incarceration by 0.37 percentage points for blacks, and 0.10 for whites. They corroborate these results using FBI Uniform Crime Reports (UCR) data for different types of offences, and conclude that the greatest impacts of graduation are associated with murder, assault, and motor vehicle theft. The authors also calibrate the social savings from crime reduction associated with completing secondary education. They show that a 1% increase in male high school graduation rates would yield $1.4 billion dollars in social benefits in 2004 dollars. Machin and Vujić (2008) (chapter 4 of this dissertation) study the relationship between crime and education using two British data sources and making use of the raisings of the school leaving age that occurred in Britain in 1947 and 1973. These data sources are twofold: individual-level data on imprisonment from the 2001 Census, as well as cohort-level panel data on offending rates from the Home Office Offenders Index Data (OID) in the period from 1984 to 2002. The main finding is that schooling significantly reduces imprisonment rates and property crime offending. Theoretical work on the relationship between human capital and crime has been done by Lochner (2004). He developed a model of crime in which human capital increases the opportunity costs of crime. The model predicts that older, more intelligent and more educated adults should commit fewer street (unskilled) crimes. It is also expected that white collar crime should decline less with age and education than unskilled crime. These predictions receive broad empirical support in self-reported data from the US.

5.3 Empirical strategy

In this chapter we use variation within pairs of twins for studying the relationship between education and crime. Obviously, the advantage of twin data is that many (unobserved) variables that twins share – like socioeconomic background and family factors – can be controlled for. Within-twin estimation has been used in several studies on the returns to schooling (see for instance Ashenfelter and Krueger, 1994 and Miller et al., 1995), and recently on the effect of parents' education on the education of their children (Behrman and Rosenzweig, 2002).

In order to get a full picture of the relationship between human capital formation and criminal behaviour, our estimation strategy consists of two steps. First, we focus on the relationship between early criminal behaviour and educational attainment. Early criminal behaviour is measured as the event of being arrested before the age of 18. It is likely that these early criminal activities occur during the time that the accumulation of human capital is still in progress because compulsory schooling laws force individuals in

Australia to attend schooling until ages 15 to 17, depending on the State of residence. For estimating the effect of early arrests on educational attainment we use the usual linear (probability) model for within-family estimation:

$$S_{ij} = \alpha + \beta A_{ij}^{17} + \gamma X_{ij} + f_j + \varepsilon_{ij} \tag{5.1}$$

where S_{ij} denotes the educational attainment of individual i in family j, A_{ij}^{17} is a dummy for being arrested before the age of 18, X_{ij} denotes a vector of covariates, f_j is an unobserved family effect common to all twins in family j, and ε_{ij} is a random error term. In this model the family fixed effect, which consists of all shared socioeconomic and genetic factors, is removed by differencing between twins. In equation (5.1), we expect that the causality primarily runs from early arrests towards educational attainment, as early arrests occur before the completion of schooling. We argue that we can largely control for reverse effects – i.e., bad school performance driving kids to start criminal activities – by including several measures of early school performance as additional controls. Moreover, we control for differences in early behaviour within pairs of twins by including an indicator of conduct disorder (see next section).

The second part of our analysis addresses the effect of human capital on crime since the age of 18, which is usually the perspective that is taken in the literature. The model that we estimate is very similar to equation (5.1):

$$C_{ij}^{18} = \alpha + BS_{ij} + \gamma X_{ij} + \delta A_{ij}^{17} + f_j + \varepsilon_{ij} \tag{5.2}$$

where C_{ij}^{18} denotes criminal activity since the age of 18. As early criminal activity is likely to be an important confounder for the estimated effect of human capital on crime, we include early arrests as an additional control. We argue that these lagged arrests can be treated as exogenous variables.

Obviously, the twin setup – together with the use of lagged information – helps us to cancel out many possible sources of endogeneity. Still, there are two important concerns in the use of within-twin estimation approach that need to be addressed in order to check the robustness of our results (Bound and Solon, 1999). First, measurement error in reported or self-reported schooling (or crime) may bias the estimates towards zero ('attenuation bias'). A solution for this problem has been introduced by Ashenfelter and Krueger (1994). They obtained two measures of schooling of a twin by asking the twins to report both on their own schooling as well as on the schooling of their sibling. The second measure of schooling can then be used as an instrument to correct for measurement error. This approach has been used in several studies (for instance Miller et al., 1995 or Behrman and Rosenzweig, 2002). In these studies, the size of the estimated effects increases after instrumenting for measurement error. This chapter follows the same approach to address any potential attenuation biases.

The second concern in within-twin models lies with respect to endogeneity bias within twin pairs. Although (identical) twins share many genes and are raised in the same social environment, they are not exactly identical. Bound and Solon (1999) show that the bias in the within-family estimator may not always be smaller than the bias in the cross-sectional estimator. This depends on the importance of the fixed family component in the unobservables. We address this possible bias by using additional controls in the within-twin models, such as conduct disorder and early arrests.

5.4 Description of data

We use data from the so-called younger cohort of twins of the Australian Twin Register (ATR). The ATR data were gathered in two surveys, in 1989-1990 and in 1996-2000. In 1980-1982 a sample of 4,262 twin pairs, born between 1964-1971, were registered with the ATR as children by their parents in response to media appeals and systematic appeals through the school system. The data were collected in two surveys among this sample of twins. In 1989-1992, when the twins were 18-25 years old, the first survey by mailed questionnaire was conducted, called Alcohol Cohort 2. The response rate of this questionnaire survey was 63%. In 1996-2000, the second survey was launched, called TWIN89. For this survey, telephone interviews were completed with 6,267 individuals, 2,805 men (889 complete and 1,027 incomplete pairs) and 3,462 women (1,215 complete and 1,032 incomplete pairs), who were 30 years old on average (from 24 to 39) at the time of the interview. The individual response rate for this telephone interview was 86%.

The surveys gathered information on the respondent's family background (parents, siblings, marital status, and children), socioeconomic status (education, employment status, and income), health behaviour (body size, smoking and drinking habits), conduct disorder, personality, feelings and attitudes. Zygosity[1] was determined by a combination of diagnostic questions plus blood grouping and genotyping.

The measures of crime used in the analysis are self-reported data on arrests and incarceration. The survey contains questions on the age of first and last arrest, the number of arrests and incarceration. The questions explicitly exclude arrests for traffic violations, drunken behaviour or drunk driving. The question on incarceration excludes time spent in jail for using drugs or alcohol.

The reliability of these self-reported data is an important issue. In criminology, the use of self-reported data is well established. Self-reporting has been the dominant technique used for measuring criminal behaviour ever since its introduction in the 1950s by Short and Nye (1957). A large literature shows that self-reported data consistently have acceptable reliability and validity. Many studies find high correlations between self-reported data and other criterion related measures of criminal frequency and arrest histories (Farrington, 1973; Hardt and Hardt-Petersen, 1977; Horney and Marshall, 1992; Huizinga and Elliott, 1986; Maddux and Desmond, 1975; Mieczkowski, 1990; Weiss et al., 1998). Thornberry and Krohn (2000) conclude that "self-reported measures of delinquency are as reliable as, if not more reliable than, most social science measures." A recent study among street-drug users recruited in 11 cities throughout the United States revealed that lifetime arrest and incarceration items demonstrated good to excellent reliability (Fisher et al., 2004). In addition, it has been shown that substance abuse factors and mental illness factors did not affect the quality and accuracy of self-reported arrests histories (Nieves et al., 2000).

Educational attainment is measured in the first survey using a seven point scale, which

[1] There are two types of twins - monozygotic (MZ) and dizygotic (DZ). The monozygotic twins are known as *identical* and are the result of the splitting of an already fertilized egg. On the other hand, dizygotic twins are know as *fraternal*, are the result of two different eggs (fertilized by two different sperm). Hence, DZ twin pairs do not have the same genetic composition like the MZ pairs, although they are more alike than pairs of siblings. The MZ twin pairs have the same genetic "make-up" because each piece of the split fertilized egg contains all and only the genetic information of the original fertilized egg (barring mutations). For more details, see for example Behrman et al. (1977).

we translate into years of education similar to Miller et al. (1995). The second survey of the younger cohort uses an eight point scale, which we also translate into years of education similar to Miller et al. (2006). We prefer to use this more recent measure, as it contains less missing values for our main estimation sample.

As covariates we use parents' education, age, and age squared. In addition, we control for conduct disorder and early school performance. Our data contains self-reported information on 21 statements that reflect behavioural problems before the age of 18 (see Section 5.A.1 in the Appendix 5.A). In the second survey the twins were asked to reflect on their experiences before the age of 18. We constructed a measure of conduct disorder by summing occurrences of these 21 statements (see Vujić et al., 2008 (chapter 6 of this dissertation)). This approach is similar to Currie and Stabile (2007) who use 6 questions to form a conduct disorder scale. Our survey also contains four questions on early school performance. Marks in primary and secondary school are measured using a three point scale: better than average, average, and below average. Respondents are also asked about the teacher's view on their school achievements: did as well as could, could have done much better, don't know. Finally, grade repetition is also measured.

In our total sample of 6267 individuals, 70 twins reported having spent time in jail and 340 twins reported having been arrested, which is 1.1% and 5.4%, respectively, of our sample. Approximately 10% of male twins and 2% of female twins reported having been arrested. A direct comparison with population statistics is complicated because of differences in reporting measures. Statistics on alleged offenders in Australia from 1995 to 2005 show that among males aged 15 to 19, approximately 9 to 13% get arrested, while among females this number is from 2 to 3% (Australian Institute of Criminology, 2007). For individuals aged 20 to 24, the rates drop to 6 to 9% for men and 2% for women. For individuals older than 24, the rates drop further to approximately 1%. It should be noted that the number of alleged offenders does not equal the number of distinct offenders during a year because police may take action against the same individual for several offences, or the individual may be processed on more than one occasion for the same offence type. In addition, we might expect that many of those arrested since the age of 20 will be recidivists. As such, a direct comparison of the arrest rates found in our data with population statistics is difficult. However, the difference between males and females seem in line with the population statistics. In addition, the total arrest rates in our sample do not seem implausibly high or low.

The sample we use for the main estimation results consists of pairs of twins with information on educational attainment and criminal participation. If this information is missing for one or both of the twins, we dropped the complete pair. In these samples, 47 twins reported having spent time in jail and 224 twins reported having been arrested. This includes 6 twin pairs (12 twins) who both report having spent time in jail and 28 twin pairs (56 individuals) who both report having been arrested. Our data contain information on the zygosity of the twins, enabling us to distinguish fraternal and identical twins. However, we only focus on the sample of all twins, pooling fraternal and identical twins together. A separate analysis on the sample of identical twins strongly reduces the sample size and especially the variation within pairs of twins on the main variable(s) of criminal behaviour. The intra-class correlation for being arrested (incarcerated) is 0.31 (0.41) for identical twins and 0.07 (0.13) for fraternal twins. Unfortunately, due to the routing of the questionnaire twins with a conduct disorder score of zero did not answer questions on criminal behaviour. As this may bias the estimates, we did some sensitivity analysis while imputing for missing values on these outcomes for twins with no childhood conduct disorder (see Section 5.7).

Table 5.1 shows sample means and proportions for educational attainment and background characteristics by criminal participation. The first two columns compare twins that spent time in jail with twins that have not been incarcerated. The last two columns compare twins that have ever been arrested with twins that have never been arrested. The sample size slightly differs between the first two columns and the last two columns because of missing values on 'incarceration' or 'having been arrested' variables. Clearly, the sample statistics show a strong association between educational attainment and participation in crime. Twins that have been incarcerated attain on average 1.6 years less education than twins that have not been incarcerated. The difference in educational attainment between those who have ever been arrested and those who have never been arrested is on average one year. Particularly remarkable are the differences in completion of senior high school, especially between those who spent time in jail and those who did not. Two thirds of those who have been incarcerated did not graduate from senior high school, compared to only one quarter of the remaining group of those who have not been incarcerated. Twins that participated in crime have lower educated parents; the difference between the columns is larger for those who spent time in jail. Male twins are more likely to be involved in criminal activity.

The bottom panel shows the statistics on conduct disorder and early school performance. The difference in conduct disorder is striking: twins who were incarcerated score approximately 5 points (2 standard deviations) higher on the indicator of conduct disorder. For twins who have ever been arrested this difference is more than two points. We also observe that twins who have been incarcerated or arrested have a higher grade repetition rate. The differences in self-reported marks in primary and secondary school seem quite modest. Moreover, the first arrest occurs much earlier for twins who have been incarcerated than for twins who have ever been arrested.

We further explore the association between human capital and crime by looking at the relationship between education and arrests. Table 5.2 shows for each schooling level the proportions for several measures of criminal participation. Criminal participation is concentrated at the two lowest schooling levels. Twins who did not complete 11 years of education are more likely to be arrested and to be incarcerated. In addition, the number of arrests is higher for those with less than 11 years of education. We also observe that many arrests of those with less than 11 years of education already take place at an early age. Moreover, their arrest rates since the age of 18 are much higher than those for twins with at least 11 or 12 years of education. Table 5.1 also makes apparent that criminal participation is fairly stable for those with at least 11 or 12 years of education. This suggests a non linear relationship between human capital and crime. Completion of senior high school (11-12 years of education) seems to be a critical boundary in this respect. Lochner and Moretti (2004) report a similar nonlinear relation between education and crime for the US. In particular, they find a steep drop in criminal participation at the level of high school graduation.

For many countries and time periods, it has well been established that crime rates

Table 5.1: Summary statistics for the main estimation samples

	Spent time in jail		Ever arrested	
	No	Yes	No	Yes
Education	11.9	10.3	12.0	11.0
	(2.4)	(2.3)	(2.4)	(2.4)
Senior high school	75.0	36.2	76.7	52.2
	(43.3)	(48.6)	(42.3)	(50.0)
Education (twin report)	11.7	10.2	11.7	10.9
	(2.3)	(2.1)	(2.3)	(2.3)
Education father	10.4	9.5	10.4	9.9
	(2.7)	(2.5)	(2.7)	(2.6)
Education mother	10.4	9.7	10.4	10.1
	(3.1)	(2.9)	(3.1)	(2.8)
Male	53.1	85.1	51.1	78.1
	(50.0)	(36.0)	(50.0)	(41.4)
Age in 1996	29.8	29.8	29.8	29.9
	(2.5)	(2.7)	(2.5)	(2.5)
Conduct disorder	3.3	8.1	3.3	5.8
	(2.5)	(3.4)	(2.4)	(3.4)
Marks primary school (1-3)	2.3	2.1	2.3	2.2
	(0.6)	(0.7)	(0.6)	(0.6)
Marks secondary school (1-3)	2.2	2.0	2.2	2.0
	(0.6)	(0.7)	(0.6)	(0.6)
Underachiever(%)	71.1	76.0	70.8	77.0
	(45.3)	(43.1)	(45.5)	(43.1)
Grade repetition(%)	18.1	31.9	18.1	24.6
	(38.5)	(47.1)	(38.2)	(0.43)
Age of first arrest	20.3	18.6		19.9
	(4.6)	(4.7)		(4.7)
Identical twin	40.7	40.4	40.7	40.6
	(49.1)	(49.6)	(49.1)	(49.2)
Estimation sample	2199	47	2028	224
Total sample	6197	70	5927	340

Table 5.2: Arrests by schooling level(%)

	Years of schooling					
	< 7	$8 - 10$	$11 - 12$	13	15	17
Ever arrested(%)	40	18.3	7.1	6.8	7.2	6.1
First arrest						
$\leq$ 15 years	20	4.0	0.6	0.5	0.7	1.1
16 years	0	1.1	0.4	0	0.3	0.6
17 years	0	2.9	0.8	1.0	0	0
18 years	0	2.3	1.1	1.9	0.7	0.5
19 years	0	0.9	0.8	1.0	0.7	0.6
Ever arrested since 18(%)	40	14.3	6.0	5.3	6.2	4.5
Number of arrests						
0	60	81.9	93.2	93.2	92.8	93.9
1	0	10.5	4.5	5.3	5.8	5.0
2	0	3.7	1.3	1.5	1.4	0.6
$\geq$ 3	40	4.0	1.0	0	0	0.6
Spent time in jail(%)	20	5.1	1.1	0.5	1.4	0.6
Number of observations	5	574	995	207	292	179

increase during the teenage years, peak around the age of twenty and decrease afterwards (Lochner 2004). This crime-age profile is well documented in criminology. Figure 5.1 shows crime-age profiles from our data based on the self-reported age of first arrest. The left panel shows an age-crime profile for individuals with less than 11 years of education. The right panel shows an age-crime profile for individuals who completed at least 11 years of education (senior high school).

The patterns in Figure 5.1 confirm the typical features of crime-age profiles from the criminology literature. That is, participation in crime increases until the age of twenty and drops afterwards. A comparison of the left and right panels suggests that individuals with less than 11 years of education start earlier with criminal activities.

5.5 The effect of early arrests on educational attainment

The strong association between education and criminal activity might be the result of early participation in crime. Early criminal involvement might be detrimental for human capital investment because of various reasons such as 'meeting the wrong friends' (building criminal capital), 'getting stigmatized', changes in motivation or aspirations. In this section we investigate the effect of early arrest on human capital accumulation by estimating linear (probability) models of early arrests on education. Table 5.3 shows estimates of the effect of early criminal participation on human capital. We use the in-

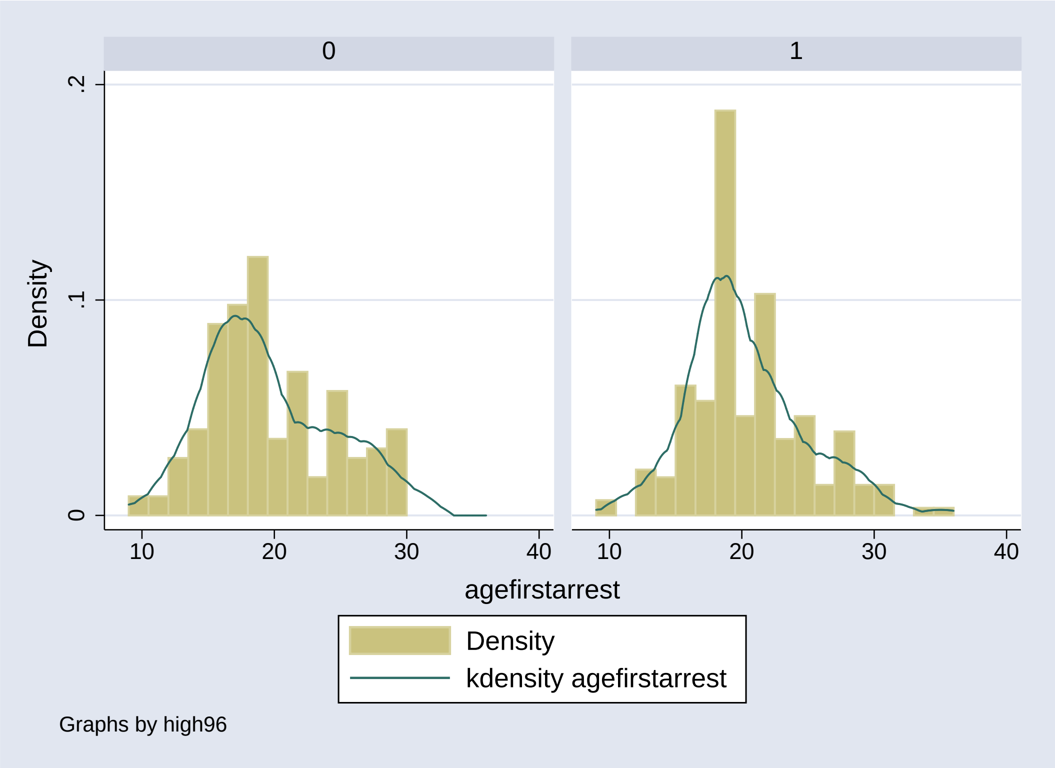

Figure 5.1: Age of first arrest by schooling level

formation on the age of first arrest as an indicator for early criminal participation and construct a dummy variable for early arrests, which equals 1 (0) if someone has (not) been arrested before the age of eighteen. Column (1) shows the OLS estimates of the effect of early arrests on educational attainment controlling for gender, age, age squared, and education of parents. Column (2) includes conduct disorder as an additional control. In column (3) additional controls for early school performance have been included: marks in primary school (1-3), marks in secondary school (1-3), grade repetition, and teachers view on underachievement. Column (4) shows the fixed effect estimates controlling for gender, column (5) controls for conduct disorder, while column (6) also includes controls for early school performance. The top panel shows the effect of early arrests on years of education, while the effects on completing senior high school are shown at the bottom of Table 5.3.

All estimates in Table 5.3 suggest that early arrests have a substantial impact on human capital accumulation. The cross sectional estimates show that those who are arrested before the age of 18 attain 1.0 to 1.5 less years of education and their probability of completing senior high school is 27 to 38 percentage points lower. The within-twin effects are smaller but remain large. Early arrests reduce educational attainment with 0.7 to 0.9 years and lower the probability of completing senior high school with 20 to 23 percentage points. Including conduct disorder reduces the effect of early arrests.[2]

[2]In case of missing values on conduct disorder, we included the value of the other twin. If both values were missing, we included the mean of the sample. In total we imputed values for 39 twins. We find

Table 5.3: Estimates of the effect of early arrests on educational attainment

	OLS (1)	OLS (2)	OLS (3)	FE (4)	FE (5)	FE (6)
Years of education						
Arrest before 18	-1.534	-1.103	-0.954	-0.856	-0.74	-0.775
	(0.235)***	(0.236)***	(0.232)***	(0.329)***	(0.329)**	(0.318)**
Conduct disorder		-0.120	-0.048		-0.084	-0.050
		(0.018)***	(0.018)***	(0.026)***	(0.026)*	(0.063)**
N	2252	2252	2252	2252	2252	2252
Twin Pairs				1126	1126	1126
Senior high school						
Arrest before 18	-0.380	-0.297	-0.270	-0.230	-0.206	-0.212
	(0.055)***	(0.056)***	(0.055)***	(0.064)***	(0.064)***	(0.063)***
Conduct disorder		-0.023	-0.014		-0.017	-0.011
		(0.004)***	(0.004)***	(0.005)***	(0.005)**	(0.012)***
N	2252	2252	2252	2252	2252	2252
Twin Pairs				1126	1126	1126

Note: All specifications control for gender. Columns (1) and (2) control for age, age squared, education of parents, columns (2) and (5) control for conduct disorder, columns (3) and (6) also control for early school performance. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

It should be noted that conduct disorder is closely related to early crime, as the 21 statements used for measuring conduct disorder include items that can be considered as criminal (see Section 5.A.1). The estimates with the third specification are quite similar to the effects of the second specification.[3] Hence, including early school performance does not affect the estimates. This indicates that, conditional on conduct disorder, the findings are determined by early criminal behaviour, rather than differences in early school performance. Another remarkable finding in Table 5.3 is the effect of conduct disorder, which is substantial for all specifications.

We further investigated the effect of the timing of the first arrest on education by constructing a second variable for early arrests. This variable measures the number of years before the age of 18 that the arrest took place (18 minus age first arrest). Table 5.4 shows the fixed effect estimates for models that include years of arrest variable and the square of this variable. Columns (1), (2), and (3) show the estimates of the effect on years of education, while columns (4), (5), and (6) show the effects on completing senior high school. We use similar controls as in Table 5.3.

Table 5.4: Estimates of the effect of the timing of the early arrest on educational attainment

	Years of education			Senior high school		
	FE	FE	FE	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)
18 minus age first arrest	-0.080	-0.652	-0.616	-0.037	-0.141	-0.126
	(0.090)	(0.252)***	(0.243)**	(0.018)**	(0.049)***	(0.048)***
(18 minus age first arrest)^2		0.099	0.097		0.018	0.016
		(0.041)**	(0.039)**		(0.008)**	(0.008)**
Conduct disorder			-0.051			-0.012
			(0.026)**			(0.005)**
N		2252	2252		2252	2252
Twin Pairs		1126	1126		1126	1126

Note: All specifications control for gender. Column (3) and (6) control for conduct disorder and early school performance. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

The estimates in Table 5.4 corroborate the previous findings. The estimation results show that the effect of early arrests also depends on the timing of the arrest, with earlier arrests being more detrimental for educational attainment. For instance, column (4) indicates that each year reduces the probability of high school completion with 4 percentage points. However, the estimates in column (5) and (6) suggest that the effect is not linear. Arrests at the age of 13, 14 or 15 are the most detrimental and reduce the probability of

similar results for the smaller sample without imputation.

[3]We imputed missing values on early school performance for 5 individuals. The results for the smaller sample without imputations are similar.

high school completion with more than 25 percentage points. Considering the fact that these arrests at age 13, 14 or 15 took place during compulsory education, these findings seems in line with our expectation that the causality runs from early arrests to human capital and not vice versa.

Summarizing, we find a large effect of early criminal behaviour on educational attainment, even when family fixed effects are taken into account. In addition, the timing of the early arrests matters, such that arrests at age 13, 14 or 15 are most detrimental for human capital accumulation.

5.6 The effect of human capital on crime

The second aspect of the strong association between education and criminal activity might be that education has an effect on crime. Investments in human capital raise the opportunity costs of crime and may also alter preferences and discount rates. Previous studies for the US and the UK find evidence for a negative effect of education on crime (Lochner and Moretti, 2004; Machin and Vujić, 2008 (chapter 4 of this dissertation)).

In this section we analyse the effect of human capital on crime. The previous section showed that reverse causality cannot be ignored, as we found substantial effects of early criminal behaviour on educational attainment. We therefore include various controls in our model that are informative on criminal behaviour before the age of 18. First, the 'early arrests' variable (arrests before the age of 18) can be used as an obvious control. Second, we can also include the 'conduct disorder' variable, which is likely to precede investments in human capital.

We use the senior high school completion variable as our main measure of human capital. Senior high school can be completed at the age of 17 or 18. This brings us the advantage that we can estimate the effect of completing senior high school on criminal activities starting from this age. The distinction between the investment in human capital and the timing of criminal activity would be less clear if we would use years of education as a measure of human capital. A second argument for using senior high school completion as a measure of human capital is that the effect of human capital on crime seems to be nonlinear (see Table 5.2).

We investigate the effect of human capital on three self-reported measures of crime: incarceration, arrests since the age of 18, and number of arrests. Unfortunately, our data do not contain information on the age of incarceration. However, statistics on incarceration in Australia show that the probability of being incarcerated before the age of 18 is very small.[4] Arrests since the age of 18 are derived from the age of the last arrest. For the number of arrests we constructed a variable which has 4 categories (0; 1; 2; 3). All individuals that reported more than three arrests were include in the last category (52 individuals reported at least three arrests, out of which 22 reported exactly three arrest). The data only contain information on the age of the first and the age of the last arrest. Hence, for the other arrests it is not clear whether they took place after the completion of

[4]The rate of non-indigenous persons aged 10 to 17 in juvenile detention between 1994 to 2003 is between 16 and 26 per 100,000 of relevant population (Charlton and McCall, 2004). This is on average approximately 0.02% of the population.

high school. Considering the evidence on reverse causality from the previous section we expect that this will give a downward bias for the estimates (more negative estimates).

Table 5.5 shows the estimates of the effect of completing senior high school on the three measures of crime, using linear probability models. The first three columns show OLS estimates, the last three columns show estimates of fixed effect models using different controls. The top panel shows the effects on the probability of incarceration, the middle panel shows the effect on the probability of being arrested since the age of 18, and the bottom panel shows the effect on the number of arrests (0-3).

The OLS estimates show that education has a negative association with all three measures of crime. This association reduces substantially when we include arrest(s) before 18 and conduct disorder. The size of the estimates further reduces when controlling for fixed family effects. All fixed effect estimates in column (4) are statistically significant. Controlling for early arrests and conduct disorder substantially reduces the size of the estimates. This confirms the earlier findings on reverse causality. Only the estimates for the effects on incarceration remain statistically significant when we include early arrests and conduct disorder measures. Completing senior high school reduces the probability of incarceration with 2.3 percentage points. Hence, the fixed effects estimates suggest that the effect of human capital on crime is only moderate.[5]

The estimates for the effect of early arrests on the three measures of crime in Table 5.5 are striking. The fixed effect estimates suggest that an early arrest increases the probability on incarceration with more than 20 percentage points and increase the probability of getting arrested since the age of 18 with 10 to 12 percentage points. In addition, the average number of arrests increases with approximately 0.2. The size of these effects are much larger than the estimated effect of completing senior high school. For instance, the estimated impact of being arrested before the age of 18 on incarceration is almost ten times higher than the estimated effect of high school completion. We also estimate the same models as in Table 5.5 with years of education instead of completing senior high school (see Table 5A.1 in the Appendix). The findings are quite similar to those in Table 5.5 and suggest a small effect of human capital on crime after controlling for early arrests and conduct disorder.

We conclude that this section provides evidence for a negative but moderate effect of human capital on crime. Completing senior high school reduces the probability of incarceration with at least 2 percentage points. Human capital also reduces the probability of being arrested since the age of 18 and the number of arrests. Including early arrests and conduct disorder measures in the models substantially reduced the effect of human capital on crime. This confirms that reverse causality is an important issue. The most remarkable findings are the large effect of early arrests on all three measures of crime. These effects are substantially larger than the estimated effects of human capital.

[5]As in the previous tables, we imputed values for 39 twins with missing data on conduct disorder. The estimation results on the smaller sample without the imputed values are similar.

Table 5.5: Estimates of the effect of high school completion on crime

Incarceration						
	OLS			Within twin estimates		
	(1)	(2)	(3)	(4)	(5)	(6)
Senior high school	-0.041	-0.020	-0.013	-0.038	-0.027	-0.023
	(0.010)***	(0.008)***	(0.008)*	(0.011)***	(0.011)***	(0.010)**
Arrest before 18		0.310	0.283		0.213	0.202
		(0.057)***	(0.056)***		(0.023)***	(0.023)***
Conduct disorder			0.008			0.008
			(0.002)***			(0.002)***
N	2246	2246	2246	2246	2246	2246
Twin pairs				1123	1123	1123
Arrested since the age of 18						
	OLS			Within twin estimates		
	(1)	(2)	(3)	(4)	(5)	(6)
Senior high school	-0.085	-0.065	-0.047	-0.037	-0.031	-0.020
	(0.016)***	(0.015)***	(0.015)***	(0.022)*	(0.022)	(0.022)
Arrest before 18		0.304	0.234		0.125	0.103
		(0.059)***	(0.057)***		(0.048)**	(0.048)**
Conduct disorder			0.021			0.018
			(0.003)***			(0.004)***
N	2252	2252	2252	2252	2252	2252
Twin pairs				1126	1126	1126
Number of arrests						
	OLS			Within twin estimates		
	(1)	(2)	(3)	(4)	(5)	(6)
Senior high school	-0.214	-0.104	-0.072	-0.108	-0.048	-0.029
	(0.034)***	(0.024)***	(0.023)***	(0.037)***	(0.034)	(0.033)
Arrest before 18		1.645	1.522		1.223	1.183
		(0.111)***	(0.106)***		(0.073)***	(0.072)***
Conduct disorder			0.037			0.033
			(0.006)***			(0.006)***
N	2250	2250	2250	2250	2250	2250
Twin pairs				1125	1125	1125

Note: All specifications control for gender. Columns (2) and (3) control for age, age squared and education of parents. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

5.7 Robustness

In this section we investigate the robustness of the findings by addressing two issues. First, we test the sensitivity of the results by imputing missing values on human capital outcomes which are due to the routing of the questionnaire. Second, we address the issue of measurement error which is likely to bias the estimates downward.

5.7.1 Missing values due to the routing of the questionnaire

Due to the routing of the questionnaire twins with a conduct disorder score of zero, which means that they reported negative on all 21 statements on conduct disorder before the age of 18, did not answer questions about arrests and incarceration. This may bias the estimates because it involves a large fraction of our sample (approximately 3000 observations). It seems likely that individuals who do not report conduct disorder will be less involved in crime than those who have a positive conduct disorder score. For instance, the arrest (incarceration) rate of those with a conduct disorder score of 3 is 7.4 (0.6) against 2.5 (0.3) for those with a conduct disorder score of 1. We checked the sensitivity of the results by imputing zeros for twins with missing values on being arrested and incarcerated. Tables 5.6 and 5.7 show the estimation results for the main models of the previous sections.

Table 5.6 shows the results for the effect of early crime on educational attainment. The estimates in Table 5.6 are somewhat smaller but quite similar to those in Table 5.3. After the imputation of the missing values for being arrested we still find a large effect of early arrests on educational attainment.

Table 5.7 shows the estimates for the effect of high school completion on crime. The pattern of findings in Table 5.7 is similar to the pattern in Table 5.5. However, the estimates of the effect of high school completion on crime in the fixed effects model that uses all controls (column (6)) becomes statistically insignificant. This suggest that the effect of educational attainment might even be smaller than indicated in Table 5.5.

We conclude that the estimates of the previous section are robust to imputing missing values of individuals with a conduct disorder score of zero. However, the estimated effect of high school completion on crime becomes statistically insignificant in models that control for early crime and conduct disorder.

5.7.2 Measurement error

A well-known concern in the literature using within-family models is measurement error (Griliches, 1979). By taking a within-family perspective, measurement error may exacerbate, which in turn is likely to bias the estimates towards zero. A solution for this problem has been proposed by Ashenfelter and Krueger (1994) in their study on the returns to schooling using data on twins. They suggest using a second independent measure of education as an instrument for educational attainment. In their study, they asked each sibling to report on both their own and their twin's schooling and used this information as an independent measures of schooling. They construct two instruments for the difference in education within twins depending on the assumptions about measurement error. Let

Table 5.6: Estimates of the effect of early arrests on educational attainment after imputations for missing values on early arrests

	OLS (1)	OLS (2)	OLS (3)	FE (4)	FE (5)	FE (6)
Years of education						
Arrest before 18	-1.597	-0.876	-0.759	-0.803	-0.624	-0.668
	(0.215)***	(0.211)***	(0.205)***	(0.289)***	(0.291)**	(0.279)**
Conduct disorder		-0.148	-0.071		-0.079	-0.038
		(0.014)***	(0.013)***		(0.018)***	(0.018)**
N	5332	5332	5332	5332	5332	5332
Twin Pairs				2666	2666	2666
Senior high school						
Arrest before 18	-0.363	-0.254	-0.225	-0.189	-0.162	-0.163
	(0.052)***	(0.051)***	(0.050)***	(0.053)***	(0.054)***	(0.053)***
Conduct disorder		-0.022	-0.014		-0.012	-0.006
		(0.003)***	(0.003)***		(0.003)***	-0.003
N	5332	5332	5332	5332	5332	5332
Twin Pairs				2666	2666	2666

Note: All specifications control for gender. Columns (1) and (2) control for age, age squared, education of parents, columns (2) and (5) control for conduct disorder, columns (3) and (6) also control for early school performance. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

Table 5.7: Estimates of the effect of senior high school completion on crime after imputing missing values on crime measures

			Incarceration			
		OLS			Within twin estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
Senior high school	-0.021	-0.010	-0.005	-0.016	-0.011	-0.008
	(0.005)***	(0.004)**	(0.004)	(0.006)***	(0.006)*	(0.006)
Arrest before 18		0.291	0.260		0.204	0.185
		(0.052)***	(0.051)***		(0.016)***	(0.016)***
Conduct disorder			0.007			0.008
			(0.001)***			(0.001)***
N	5326	5326	5326	5326	5326	5326
Twin pairs				2663	2663	2663
			Arrested since the age of 18			
		OLS			Within twin estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
Senior high school	-0.054	-0.042	-0.025	-0.023	-0.020	-0.011
	(0.009)***	(0.009)***	(0.008)***	(0.012)*	(0.012)*	(0.012)
Arrest before 18		0.315	0.218		0.14	0.094
		(0.054)***	(0.052)***		(0.033)***	(0.033)***
Conduct disorder			0.021			0.021
			(0.002)***			(0.002)***
N	5332	5332	5332	5332	5332	5332
Twin pairs				2666	2666	2666
			Number of arrests			
		OLS			Within twin estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
Senior high school	-0.127	-0.065	-0.037	-0.058	-0.027	-0.013
	(0.020)***	(0.013)***	(0.013)***	(0.020)***	(0.018)	(0.018)
Arrest before 18		1.651	1.491		1.261	1.184
		(0.101)***	(0.096)***		(0.050)***	(0.049)***
Conduct disorder			0.035			0.036
			(0.004)***			(0.003)***
N	5330	5330	5330	5330	5330	5330
Twin pairs				2665	2665	2665

Note: All specifications control for gender. Columns (2) and (3) control for age, age squared and education of parents. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

S_1^1 refer to the self-reported education level of the first twin, S_1^2 to the sibling-reported education level of the first twin, S_2^2 to the self-reported education level of the second twin and S_2^1 to the sibling-reported education level of the second twin. The first instrument uses the difference in the twins' report on the schooling of their sibling as an instrument for the difference in the report on own schooling. Hence, $S_1^1 - S_2^2$ is instrumented with $S_1^2 - S_2^1$. The second instrument assumes that the measurement error of respondent's report on own schooling and the schooling of their sibling is correlated. In the estimation the difference in the reports of twin A about own schooling and the sibling's schooling is instrumented with the difference in the reports of twin B on the sibling's schooling and own schooling. Hence, $S_1^1 - S_2^1$ is instrumented with $S_1^2 - S_2^2$.

In our study we can follow this approach in the models that estimate the effect of education on crime because our data include the same questions on the sibling's schooling. The correlation between the self-reported level of education and the sibling-reported education level, which indicates the reliability ratio, is 0.80. For high school completion this correlation is 0.63. Unfortunately, our data do not contain sibling's reports on criminal behaviour. Therefore, we can only use this approach for the models that investigate the effect of education on crime. It should be noted that this approach produces consistent estimates when the measurement error is classical. However, since our main variable (senior high school completion) is a binary indicator, the measurement error is non-classical. It has been shown that the IV estimate will then be upward biased (Aigner, 1973; Kane et al., 1999). The within-family estimate from the previous analysis will then provide a lower bound and the IV estimate an upper bound of the true (negative) effect.

Table 5.8 shows the IV estimates for the effect of high school completion on the three measures of crime. Columns (1), (3), and (5) show the estimation results for the first instrument described above. Columns (2), (4), and (6) show the results for the second instrument. All specifications use early arrest, conduct disorder, and gender as controls. First stage results confirm validity of instruments.

The estimates in Table 5.8 suggest that measurement error in education might be important. All estimates increase and most estimates are statistically significant. The estimates with the first instrument are very large but also have large standard errors. The estimates with the second instruments are also larger than the estimates in Table 5.5 but more precise. These results suggest that the findings in Table 5.5 might underestimate the true effect of human capital on crime. We find a similar pattern when using years of education instead of completion of senior high school. However, the estimates are smaller (see Table 5A.2 in the Appendix). In addition, we re-estimated the models from Table 5.8 after imputing the missing values for individuals with a conduct disorder score of zero (see Table 5A.3 in the Appendix). The size of the estimates is smaller after the imputations, but the pattern of findings remains similar.

The findings from this sections suggest that the previous estimates might be downward biased because of measurement error. The IV estimates suggest larger effects of education on crime but the estimates are less precise.

Table 5.8: IV estimates of the effect of senior high school completion on crime

A. Crime Equations	Jail		Arrest since 18		Number of arrests	
	IV1	IV2	IV1	IV2	IV1	IV2
	(1)	(2)	(3)	(4)	(5)	(6)
Senior high school	-0.199	-0.060	-0.231	-0.086	-0.401	-0.105
	(0.117)*	(0.024)**	(0.228)	(0.050)*	(0.348)	(0.075)
N	2243	2243	2249	2249	2247	2247
Twin pairs	1123	1123	1126	1126	1125	1125
B. First Stage Equations						
IV1 / IV2	0.099	0.509	0.099	0.506	0.099	0.506
	(0.029)***	(0.024)***	(0.029)***	(0.024)***	0.029)***	(0.024)***
N	2243	2243	2249	2249	2247	2247
Twin pairs	1123	1123	1126	1126	1125	1125
F-test on instruments	11.48	434.34	11.59	432.63	11.58	432.23
p-value	0.000	0.000	0.000	0.000	0.000	0.000

Note: All specifications control for gender, early arrest and conduct disorder. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

5.8 Decomposing crime–education relationship

The two main findings from the previous sections are that early criminal behaviour is detrimental to investment in human capital and that human capital has a negative effect on crime. In this section we try to assess the importance of these two effects for the association between crime and education. We estimated within-twin models of the effect of education on 'ever been arrested' and inspected how the estimated effect of education changes after including early crime and conduct disorder. Including 'early arrests' in the estimation controls for the effect of early criminal behaviour on educational attainment and 'explains' all arrests before the age of 18, leaving only crime since 18 to be explained. Table 5.9 shows the estimation results using years of education or high school completion as explanatory variables.

The estimates in the first column show that one year of education is associated with a reduction of the probability of being arrested with 1 percentage point. Completion of high school is associated with a reduction of the probability of being arrested with 6.7 percentage points. The estimates of the effect of human capital reduce dramatically after the inclusion of 'early arrest' (column (2)). The estimated effect of one year of education reduces to 0.4 percentage points and the estimated effect of high school completion to 2.6 percentage points. Including conduct disorder further reduces the estimated effects to 0.2 and 1.7 percentage points (column (3)). In other words, controlling for early arrests and early behaviour problems reduces the estimated effect of human capital on crime to less than a quarter of the previously estimated association. From this, we conclude that early criminal behaviour explains most of the association between human capital and crime.

Table 5.9: Fixed effect estimates of the effect of education on the probability of being arrested

	(1)	(2)	(3)
Years of education	-0.010	-0.004	-0.003
	(0.005)**	(0.004)	(0.004)
N	2252	2252	2252
Twin pairs	1126	1126	1126
Senior high school	-0.067	-0.026	-0.018
	(0.024)***	(0.021)	(0.021)
N	2252	2252	2252
Twin pairs	1126	1126	1126
Controls			
Early arrest	No	Yes	Yes
Conduct disorder	No	No	Yes

Note: All specifications control for gender. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

5.9 Conclusions and discussion

This chapter aimed to disentangle the strong association between human capital and crime by investigating whether crime reduces investment in human capital or whether education reduces criminal activity. Heretofore, we exploit two aspects of the Australian survey data that we use. First, as the data are obtained from twins, we are able to control for many unobserved characteristics affecting both criminal behaviour and the schooling decisions. Second, as criminal behaviour is measured over different periods of time – prior to and after senior high school completion – we can address the causality between crime and education as well. As early criminal behaviour may affect human capital formation, and human capital may influence criminal behaviour in later stages of life, we follow a two step analysis.

First, we address the effects of early criminal behaviour on educational attainment. The estimates suggest that early criminal behaviour is detrimental to investment in human capital. Within pairs of twins we find that early arrests (before the age of 18) reduce educational attainment with 0.7 to 0.9 years and lower the probability of completing senior high school with 20 to 23 percentage points. In addition, the timing of the early arrest matters, such that arrests at age 13, 14 or 15 are most detrimental for educational attainment. These estimates are found after controlling for conduct disorder and early school performance.

Second, we focus on the effect of human capital on crime. As early criminal activity might be an important confounder, we control for early arrests. The estimates suggest that human capital has a negative effect on crime. Completing senior high school reduces the probability of incarceration with 2 to 3 percentage points. We find similar but statistically insignificant effects on the probability of being arrested since the age of 18 and on the number of arrests. The size of these estimates might be downward biased because of measurement error in schooling. IV estimates using a second independent measure of schooling suggest that the effect of human capital might be larger. In comparison, Lochner and Moretti (2004) report IV estimates of the effect of high school completion on imprisonment of 8 percentage points for blacks and 0.9 for whites.

When combining these findings, it seems that the causality between human capital and crime runs in both directions. Still, the impact of early criminal behaviour on human capital formation dominates the impact of human capital formation on future criminal behaviour. Controlling for early arrests and early behaviour problems reduces the estimated effect of human capital on crime to less than a quarter of the previously estimated association. Therefore, we conclude that early criminal behaviour explains most of the association between human capital and crime.

The strong detrimental effects of early criminal behaviour become also transparent if we consider the estimated effects of early arrests on all three measures of crime. Early arrests increase the probability of incarceration with 20 percentage points and the probability of being arrested since the age of 18 with 10 percentage points. These effects are much larger than the estimated effects of human capital. For instance, the estimated effect of being arrested before the age of 18 on incarceration is almost ten times higher than the estimated effect of completing high school.

In line with previous studies (Lochner and Moretti, 2004; Machin and Vujić, 2008 (chapter 4 of this dissertation)), our findings suggest that policies that succeed in raising investment in human capital might reduce crime. However, the (direct) returns to polices that succeed in preventing early criminal behaviour might be much larger. The estimated effects of early criminal behaviour and conduct disorder stress the importance of the early stages of life for preventing crime. Programs that keep children on 'the right track' not only may yield high private returns but also may yield high social returns through their impact on crime reduction. Studies on the effects of effective early schooling programs in the US show that these program have large social returns mainly through their impact on preventing crime (Carneiro and Heckman, 2003).

Our main conclusion is that the strong association between human capital and crime is mainly driven by the effect of early criminal behaviour on educational attainment. This finding, based on within-twin estimation approach, confirms one of the main conclusions from a synthesis of the literature on the causes of crime: 'We must rivet our attention on the earliest stages of the life cycle, for after all is said and done, the most serious offenders are boys who begin their criminal careers at a very early age," (Wilson and Herrnstein, 1985; cited in DiIulio, 1996).

5.A Appendix

5.A.1 Variable definitions

Variable	Question
misbehaved	**L3** Did you *frequently* get into a lot of trouble with the teacher or principal for misbehaving in school? (primary or secondary school)
wagged school	**L4** Before age 18, did you ever wag school for an entire day at least twice in 1 year?
suspended/expelled	**L5** Were you ever suspended or expelled from school?
stay out late	**L6** As a child or a teenager, did you *often* stay out *much later* than you were supposed to?
sneak out at night	**L6A** Did you *often* sneak out of the house at night?
run away overnight	**L6C** Before age 18, did you ever run away from home *overnight*?
lied, used false name	**L7** Before 18, did you ever tell *a lot* of lies or use a false name or alias?
outsmarted, conned others	**L7B** Before age of 18, was there ever a period when you often outsmarted others and ŞconnedŤ them?
stole from home or family	**L8** Before age 18, did you steal money or things from your home or family more than once? If yes, did you only steal things of trivial value, like loose change or things like that?
shoplifted	**L8A** Before age 18, did you steal or shoplift from shops or other people (without their knowing) more than once? If yes, did you only steal things of trivial value like comics or lollies?
forged signature	**L8B** Before age 18, did you forge anyoneŠs signature on a cheque or credit card more than once?
damaged property	**L9** Have you ever damaged someoneŠs property *on purpose*?
started physical fights	**L10** Before age 18, did you start physical fights (with persons *other than* your brothers or sisters) 3 or more times?
used a weapon	**L11** Before age 18, did you ever use a weapon like a bat, brick, broken bottle, gun or a knife (other than in combat, when hunting, or as part of your job) to threaten or harm someone?
physically injured someone	**L12** Before age 18, (other than fighting or using a weapon) did you ever physically injure anyone on purpose?
bullied others	**L13** Before age 18, were you *often* a bully, deliberately hurting or being mean to others?
mean to animals	**L14** Before age 18, were you ever mean to animals including pets or did you hurt animals on purpose?
lighted fires	**L15** Before 18, did you ever deliberately light any fires you were not supposed to?
broke into someoneŠs car/house	**L16** Before 18, did you ever break into someoneŠs car or house or anywhere else (not because you were locked out)?
forcefully stole money or property	**L17** Before age 18, did you ever take money or property from someone else by threatening them or using force, like snatching a purse or robbing them?
forced someone into sexual activity	**L20** Before age 18, did you ever force anyone into intercourse or any other form of sexual activity?

Note: Conduct disorder statements from the TWIN89 questionnaire.

5.A.2 Robustness checks tables

Table 5A.1: Estimates of the effect of years of education completed on crime

			Incarceration			
		OLS			Within twin estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
Years of education	-0.006	-0.003	-0.001	-0.004	-0.002	-0.001
	(0.002)***	(0.001)**	(0.001)	(0.002)*	(0.002)	(0.002)
Arrest before 18		0.313	0.284		0.218	0.205
		(0.057)***	(0.056)***		(0.023)***	(0.023)***
Conduct disorder			0.009			0.009
			(0.002)***			(0.002)***
N	2246	2246	2246	2246	2246	2246
Twin pairs				1123	1123	1123
			Arrest since 18			
		OLS			Within twin estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
Years of education	-0.012	-0.009	-0.006	-0.006	-0.005	-0.003
	(0.003)***	(0.003)***	(0.003)**	(0.004)	(0.004)	(0.004)
Arrest before 18		0.314	0.242		0.128	0.104
		(0.059)***	(0.057)***		(0.048)***	(0.048)**
Conduct disorder			0.021			0.018
			(0.003)***			(0.004)***
N	2252	2252	2252	2252	2252	2252
Twin pairs				1126	1126	1126
			Number of arrests			
		OLS			Within twin estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
Years of education	-0.032	-0.016	-0.010	-0.019	-0.011	-0.007
	(0.006)***	(0.004)***	(0.004)**	(0.007)***	(0.007)	(0.007)
Arrest before 18		1.660	1.532		1.225	1.184
		(0.111)***	(0.106)***		(0.073)***	(0.072)***
Conduct disorder			0.038			0.032
			(0.006)***			(0.006)***
N	2250	2250	2250	2250	2250	2250
Twin pairs				1125	1125	1125

Note: All specifications control for gender. Columns (2) and (3) control for age, age squared and education of parents. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

Table 5A.2: IV estimates of the effect of years of education on crime

	Jail		Arrest since 18		Number of arrests	
	IV1	IV2	IV1	IV2	IV1	IV2
	(1)	(2)	(3)	(4)	(5)	(6)
Years of education	-0.005	-0.003	-0.006	-0.007	-0.008	-0.008
	(0.005)	(0.003)	(0.010)	(0.007)	(0.015)	(0.011)
N	2243	2243	2249	2249	2247	2247
Twin pairs	1123	1123	1126	1126	1125	1125

Note: All specifications control for gender, early arrest and conduct disorder. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

Table 5A.3: IV estimates of the effect of senior high school completion on crime after imputing missing values on crime measures

	Jail		Arrest since 18		Number of arrests	
	IV1	IV2	IV1	IV2	IV1	IV2
	(1)	(2)	(3)	(4)	(5)	(6)
Senior high school	-0.106	-0.030	-0.117	-0.061	-0.283	-0.094
	(0.066)	(0.013)**	(0.130)	(0.026)**	(0.202)	(0.039)**
N	5322	5322	5328	5328	5326	5326
Twin pairs	2663	2663	2666	2666	2665	2665

Note: All specifications control for gender, early arrest and conduct disorder. Standard errors in brackets. ***/**/* significant at 1%/5%/10%-level.

Chapter 6

The Effect of Childhood Conduct Disorder on Human Capital: Evidence from a Cohort of Young Australian Twins

This chapter is based on work done at the CPB during 2006–2008, under the supervision of Dr. Dinand Webbink and Dr. Pierre Koning.

6.1 Introduction

Many children have mental health problems which hinder their normal development and functioning. Anxiety, conduct, attention, and depressive disorders are the most common. For example, up to 500,000 (14%) of Australian children between the ages of 4 and 18 years have significant mental health problems (Australian Centre for Community Child Health, 2006); up to 50% of 'hard to manage' pre-school children are likely to have continuing hardships requiring professional help; approximately one in five children and adolescents in the US may have a mental health disorder (Currie and Stabile, 2006). Despite these large numbers little is known about the longer term effects of these mental health problems of children. Currie and Stabile (2007) note that most studies "assume that early mental health problem will have negative effects and focus on the efficacy of specific interventions." In the economic literature several recent studies investigate the longer term effects of mental health problems on human capital. Currie and Stabile (2006) and Fletcher and Wolfe (2008) investigate the effects of attention deficit hyperactivity disorder (ADHD) in the US and Canada. Currie and Stabile (2007) also investigate the effects of depression and conduct disorder. Le et al. (2005) analyse the effects of conduct disorder on early school leaving and labour market outcomes in Australia. These papers find that mental disorders, and especially ADHD, have large negative effects on human capital accumulation.

This chapter analyses the longer term effects of childhood conduct disorder on human capital accumulation and violent and criminal behaviour later in life using data of Australian twins. If someone shows "a repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate societal norms or rules are violated," then he/she falls into a category of individuals with a conduct disorder problem. Conduct disorder is known as a 'disruptive behaviour disorder' because of its impact on children and their families, neighbours, and schools, and is largely associated with delinquent or criminal activity. We measure conduct disorder using diagnostic criteria from the American Psychiatric Association (APA, 1994). In line with Currie and Stabile (2007) we estimate effects on positive human capital, including measures such as grade repetition, marks in primary and high school, and educational attainment, and on 'negative' human capital, including measures such as being arrested, spent time in jail, and physically attacking others.

The two main challenges in estimating the longer term effects of childhood conduct disorder on human capital are the measurement of conduct disorder and omitted variable bias. This chapter contributes to the economic literature by addressing these two main challenges. First, it is easy to know when a child has a fever but a child's mental health problem is harder to identify. Typically, mental health problems in children are diagnosed by asking a child's parents and teachers a series of questions about their behaviours. For instance, Currie and Stabile (2007) use 6 questions to form a conduct disorder scale. In this study we can employ a much richer set of conduct disorder indicators. We use self reports of adult twins on 21 statements that follow the definition of conduct disorder according to the APA criteria. In addition, the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) emphasizes that there should be at least three specific conduct disorder behaviours present within the same 12 month period to make the diagnosis of conduct disorder (American Psychiatric Association, 1994). Our data include self reports on a question that exactly matches this definition. Moreover, we have information whether the twin saw a doctor, psychologist or other professional for these behaviours. We use these information for constructing three measures of conduct disorder. Another contribution of this study is that we also have information on the age at which the conduct disorder behaviours occurred. We use this information for investigating the effects of the timing of the problem behaviours on human capital.

Second, estimates of the effect of conduct disorder on human capital might be confounded by unobserved differences between children and their families. For example, in Australia the incidence of mental health problems is even higher in disadvantaged children, such as Aboriginal children (24%), children residing in 'out of home care' (55-60%) and children with a disability, who are up to four times more likely to have mental health problems than children without a disability (Australian Centre for Community Child Health, 2006). Parental substance abuse, marital conflict, psychiatric illness, and child abuse and neglect have been identified as risk factors for conduct disorder (Searight et al., 2001). Previous studies on the effects of childhood mental health problems mainly used cross sectional and within-sibling estimations (Currie and Stabile, 2006, 2007; Fletcher and Wolfe, 2008). In this chapter we estimate within-family models using data on (identical) twins. We are aware of only one previous study (Le et al., 2005) that also uses data on twins in the context of conduct disorder effects on human capital accumulation. The advantage of using twins instead of siblings is that the family circumstances for twins will typically be more similar than with siblings. More importantly, identical twins are genetically identical, whereas siblings, on average, share only half of their genetic endowments. Therefore, using data on twins, in particular genetically identical twins, may reduce the bias caused by heterogeneity within families.

We find large deleterious effects of conduct disorder on positive and negative human capital, even within pairs of identical twins. Conduct disorder increases the probability of grade retention and not completing high school. In addition, conduct disorder behaviours measured before the age of 18 have a strong effect on violent and criminal behaviour since the age of 18. For instance, within pairs of identical twins we find that conduct disorder reduces the probability of high school graduation with 4 to 13 percentage points and increases the probability of being arrested with 7 to 16 percentage points. Another important finding is that the effect of conduct disorder on human capital is more deleterious if these behaviours occur earlier in life. Various robustness checks, for instance sensitivity analyses that address the issues of measurement error and spill-over effects within pairs of twins, suggest that our main estimates may be lower bounds of the true effect of conduct disorder on human capital. We conclude that childhood mental health problems have high human and financial costs for families and society at large. Effective treatment early in life may yield high returns.

The remainder of this chapter is organised as follows. The next section reviews previous studies. Section 6.3 describes the data, while the methodology is explained in Section 6.4. Section 6.5 shows the main estimation results. Robustness checks are presented in Section 6.6. Section 6.7 concludes and discusses some policy implications.

6.2 Previous studies

This chapter examines the relationship between conduct disorder problems early in life and future human capital accumulation. We focus on conduct disorder problems as a measure of non-cognitive traits, as opposed to IQ tests that measure intelligence or cognitive ability. The line between the cognitive and non-cognitive traits is not very clear and has not been consistently defined in either the psychology or economics literature (Borghans et al., 2008). Three strands of the previous literature are related to our study. The first one, which is probably most similar to our work, consists of health economics papers that look at early childhood health problems and their effects on school performance and educational attainment (Currie and Stabile, 2006, 2007; Fletcher and Wolfe, 2008; Slade and Wissow, 2007). The second strand of literature examines the relationship between child development, school attainment, and labour market outcomes (Gregg and Machin, 2000; Le et al., 2005). The third group of papers focuses on the importance of non-cognitive skills on labour market outcomes and social behaviour (Borghans et al., 2008; Carneiro and Heckman, 2003; Heckman et al., 2006; Blanden et al., 2006; Heckman and Masterov, 2007).

First, Currie and Stabile (2006) examine the relationship between hyperactivity (ADHD) symptoms and short-term human capital outcomes (test scores, grade repetition, special education, and delinquency) using longitudinal data. The measurement of ADHD was based on questions asked to parents and/or teachers of US children (aged 4-14) or Canadian children (aged 4-11). Using ordinary least squares and within sibling estimation they find large negative effects on test scores and schooling attainment. In a follow up study, Currie and Stabile (2007) also pay attention to depression and conduct disorder (antisocial behaviour/aggression). Conduct disorder has been measured using 6 questions. They find that children with mental health problems, especially with ADHD, suffer large negative consequences in terms of future human capital outcomes. For the US, they find that conduct disorder has negative effects on various human capital outcomes, while for Canada they only find negative effects on the probability that 16-19 year old youths are in school (fixed effect estimates for the other outcomes suggest negative effects but are statistically insignificant). In this ball park of literature fall recent papers by Fletcher and Wolfe (2008) and Slade and Wissow (2007). Fletcher and Wolfe (2008) estimate the effect of ADHD on human capital. Measuring ADHD using a series of survey questions, they look at the long-term relationship between childhood symptoms of ADHD and human capital accumulation (grade point average, secondary school indicators, years of education, and probability of attending college). Their OLS results imply that children with ADHD face longer-term educational disadvantages. Once family fixed effects are controlled for, standard errors become larger, thus rendering insignificant estimation results. However, they find evidence for spill-over effects within families. Using the same US data on adolescent health, Slade and Wissow (2007) look at the connection between childhood maltreatment, which causes emotional and behavioural problems throughout childhood, and academic performance in middle and high school. Using within family estimation approach, they find that more intense childhood maltreatment reduces grade point averages and leads to problems completing homework assignments.

Second, Gregg and Machin (2000) examine the relationship between child development and labour market outcomes. Using sequential modelling and UK data, they find that behavioural problems at age 7 are associated with poorer educational outcomes at age 16, which in turn is associated with poor labour market outcomes at ages 23 and 33. Further, the children of parents who grew up in the socially disadvantaged situation during their own childhood, have lower early age cognitive abilities, suggesting an important cross-generational link. Le et al. (2005) investigate the effect of childhood conduct disorder on early school leaving and labour market outcomes, such as employment and earnings, in Australia. They find that individuals who experienced conduct disorder problems are more likely to leave school early, have poorer employment prospects and lower earnings. However, the within-twin estimates of the effect of conduct disorder on early school leaving conducted on the sample of identical twins are statistically not significant. They therefore conclude that genetic factors are responsible for the positive link between conduct disorder and early school leaving. In this chapter we use the same data as in Le et al. (2005), but focus on different outcomes, with one exception, and use three measures of conduct disorder, instead of one measure (see next section).

Finally, several recent papers stress the importance of non-cognitive skills for labour market outcomes and social behaviour (Borghans et al., 2008; Carneiro and Heckman, 2003; Heckman et al., 2006; Blanden et al., 2006; Heckman and Masterov, 2007). Borghans et al. (2008) examine the line between economics and psychology and ask whether "economics would benefit from incorporating the findings of personality psychology to better predict and understand economic outcomes." They present extensive evidence on the predictive power of personality traits and consider whether personality traits can be changed by intervention. Heckman et al. (2006) present evidence that both cognitive and non-cognitive skills play a dominant role in explaining personal achievement. They conclude that non-cognitive skills are as equally important as cognitive skills in determining social and economic success in life. Measures of non-cognitive skills are likely to capture some aspects of mental health as well as innate ability traits. Blanden et al. (2006) examine whether rising returns to non-cognitive skills can explain growing income inequality. Analysing 1958 and 1970 British birth cohort data sets, they include characteristics such as "hyper" and "anxious" as well as measures such as "self esteem" and "extrovert" as measures of non-cognitive skills and find that rising returns to positive mental characteristics do indeed account for some of the increase in inequality between the two cohorts. Carneiro and Heckman (2003) review the early childhood and adolescent years intervention programmes in the US that proved to be effective "in reducing criminal activity, promoting social skills of young adults, and integrating disadvantaged children into the mainstream society."[1] They show that both cognitive and non-cognitive abilities affect schooling and economic success. They also show that socio-economic differences in cognitive and non-cognitive abilities appear early in life and widen over the life cycle of the child. Hence, intervention programmes aimed at adolescents coming from bad family environments can positively affect learning and subsequent employment and earnings, but cannot reverse the early childhood damage. Heckman and Masterov (2007) go one step further and give a productivity argument for investing in young children from disadvantaged environments. Evidence shows that these children are more likely to commit crime, have out-of-wedlock births, and drop out of school. Early interventions that alleviate the effects of adverse environments can be beneficial not only to children themselves, but also to their own children, and society at large.

6.3 Data

In this study we analyse data from the so-called younger cohort of twins of the Australian Twin Register (ATR). The data were gathered in two surveys, in 1989-1990 and in 1996-2000. In 1980-1982 a sample of 4,262 twin pairs, born between 1964-1971, were registered with the ATR as children by their parents in response to media appeals and systematic appeals through the school system. In 1989-1992, when the twins were 18-25 years old, the first survey by mailed questionnaire was conducted, called Alcohol Cohort 2. The response rate of this questionnaire survey was 63%. In 1996-2000, the second survey was launched, called TWIN89. Telephone interviews were completed with 6,267 individuals, 2,805 men (889 complete and 1,027 incomplete pairs) and 3,462 women (1,215 complete and 1,032 incomplete pairs), who were 30 years old on average (range from 24 to 39) at the time of the interview. The individual response rate for this telephone interview was 86%. The surveys gathered information on the respondent's family background (parents, siblings, marital status, and children), socioeconomic status (education, employment status, and

[1] Early intervention programmes in the US, like the Perry Pre-School Programme (PPP), the Syracuse Programme (SP) or the Head Start Programme (HSP), have shown to be successful, see for example Schweinhart et al. (1993), Donohue and Siegelman (1998), Lally et al. (1988), and Garces, Thomas, and Currie (2002).

income), health behaviour (body size, smoking and drinking habits), personality, feelings and attitudes. Zygosity[2] was determined by a combination of diagnostic questions plus blood grouping and genotyping.

6.3.1 Conduct disorder

The measurement of conduct disorder is crucial for our analysis. According to the American Psychiatric Association (APA, 1994), a person with conduct disorder problems shows "a repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate societal norms or rules are violated, as manifested by the presence of three (or more) of the following criteria in the past 12 months." For instance criteria like: often initiated physical fights; has deliberately destroyed others' property; has broken into someone else's house, building, or a car; has often been truant from school, etc. (for a full definition, see Section 6.A.2 in the Appendix 6.A). Our data contains self-reported information on 21 statements that reflect behavioural problems before the age of 18 (for a full set of questions, see Section 6.A.1 in the Appendix 6.A). The information on conduct disorder is based on the second survey among adult twins (ages 24-39). Twins were asked to reflect on their experiences before the age of 18. As our first measure of conduct disorder we created a conduct disorder score based on these 21 statements. We first constructed a conduct disorder index by using factor analysis. However, the main factor appeared to be highly correlated with a conduct disorder score based on summing occurrences of these 21 statements.[3] As this conduct disorder score is more comparable to measures used in previous studies, for instance the 'antisocial/aggression score' from Currie and Stabile (2007), we prefer to use this measure, which we call the 'Conduct disorder score'. Le et al. (2005) used 19 statements for their measure of conduct disorder.[4] Our second measure of conduct disorder is based on the question 'Did you do at least 3 of these things within the same 12-month period?' Twins who responded 'yes' were coded as 1, twins who responded no or twins with a conduct disorder score of 0, 1 or 2 were coded as 0. As this measure perfectly matches the APA definition of conduct disorder, we will use it as our main measure in the analysis. We call this measure the 'APA Definition'. Our survey also asked 'How old were you the first time you did at least 3 of these things within the same 12-month period?' We use this information for investigating the effect of the timing of conduct disorder. Our third measure of conduct disorder is based on the question 'Did you ever see a doctor, psychologist, social worker or any other professional like that because of these behaviours?'. Twins who responded 'yes' were coded as 1, twins who responded 'no' or twins who were coded as 0 on the APA definition were coded as

[2] There are two types of twins - monozygotic (MZ) and dizygotic (DZ). The MZ twins are known as *identical* and are the result of the splitting of an already fertilized egg. On the other hand, DZ twins are know as *fraternal*, are the result of two different eggs (fertilized by two different sperm). Hence, DZ twin pairs do not have the same genetic composition like the MZ pairs, although they are more alike than pairs of siblings. The MZ twin pairs have the same genetic make-up because each piece of the split fertilized egg contains all and only the genetic information of the original fertilized egg (barring mutations). For more details, see for example Behrman et al. (1977).

[3] The correlation between the conduct disorder index based on factor analysis and the conduct disorder score is 0.994.

[4] We also included 'misbehaved' and 'suspended/expelled'.

0 on this third measure. We call this measure 'Professional help'. Seeking professional help may signal that conduct disorders hinder normal development and human capital accumulation. Obviously, this measure is endogenous as people choose or are advised to seek professional help. Le et al. (2005) did not use the second and third measures of conduct disorder.

6.3.2 Positive and negative human capital

In line with Currie and Stabile (2007), we distinguish between positive and negative human capital. Positive human capital is the type of human capital that is generally accumulated through schooling. Negative human capital is related with criminal, antisocial and violent activities. Our data contain five measures of positive human capital: grade repetition, marks in primary and high school, more than three college or university dropouts and educational attainment. Marks in primary and high school are measured with a three point scale (below average, average, better than average). Educational attainment is measured using an eight point scale: less than 7 years of schooling; 8-10 years of schooling; 8-10 years of schooling and apprenticeship or diploma; 11-12 years of schooling; apprenticeship, diploma, certificate; technical or teachers' college; university, first degree; university, postgraduate degree. These eight categories are recorded as 5, 9, 9, 11.5, 11.5, 13, 15 and 17 years of education, respectively (Miller et al., 2006). From this variable we created a dummy for high school graduation (at least 11.5 years of education completed), which is our sixth measure of positive human capital. We also use 6 measures of 'negative' human capital: 'more than three job quits'; 'lying, using false names or alias, or conning others'; 'physically attacking others'; 'failing to pay debts'; 'being arrested since age 18'; 'spent time in jail'). All these negative human capital variables are dichotomous variables. Unfortunately, due to the routing of the questionnaire twins with a conduct disorder score of zero did not answer several questions on positive/negative human capital. As this may bias the estimates downward we checked the sensitivity of the results by imputing zeros on these outcomes for twins with no childhood conduct disorder (see Section 6.6). In the analysis we use as covariates: mother's and father's education, age, age squared, gender, and birth weight.

The main variables in our analysis (conduct disorder, positive and negative human capital) are all based on self-reports. The reliability of these self-reported data is an important issue. In this chapter we use three different measures of conduct disorder and one of our measures is based on 21 statements. We expect that the use of such a rich set of indicators improves the reliability of the data. In criminology, a large literature shows that self-reported data consistently have acceptable reliability and validity. Many studies find high correlations between self-reported data and other criterion related measures of criminal frequency and arrest histories (Farrington, 1973; Hardt and Hardt-Petersen, 1977; Horney and Marshall, 1992; Huizinga and Elliott, 1986; Maddux and Desmond, 1975; Mieczkowski, 1990; Weiss et al., 1998). Thornberry and Krohn (2000) conclude that "self-reported measures of delinquency are as reliable as, if not more reliable than, most social science measures." Further, Le et al. (2005) who use the same Australian data, report that the twin data sample is representative of comparable segments of the general population.

6.3.3 Descriptive statistics

Table 6.1 reports means of all measures for pairs of twins with non-missing information on conduct disorder symptoms. Columns 1 and 3 show means for complete pairs of all and identical twins respectively. Columns 2 and 4 show the number of twins with a within-family difference in the variable in question for the samples of all and identical twins respectively. The top panel shows that the average conduct disorder score is nearly 2. Currie and Stabile (2007) report averages of nearly 5 for the US and 1.5 for Canada using a scale from 0 to 16. Hence, our Australian sample seems more comparable to the Canadian sample than to the US sample. The second row of Table 6.1 shows that 13% of our sample would be diagnosed as having conduct disorder according to the APA definition, for men and women this is respectively 20% and 8%. This is roughly in line with Searight et al. (2001) who report that approximately 6% to 16% of boys and 2% to 9% of girls meet the diagnostic criteria for conduct disorder. Additionally, this is also in line with Australian Centre for Community Child Health (2006), which reports that 14% of Australian children between the ages of 4 and 18 years have significant mental health problems. The fraction of our sample that saw professional help is much smaller, approximately 3%.

The number of twins with a within-family difference on one of the three measures of conduct disorder is much larger in the sample of all twins than in the sample of identical twins (columns 2 and 4). This illustrates that genetic factors, that are exactly the same within pairs of identical twins, are important for the development of conduct disorder. The average (standard deviation) of the within twin differences for the conduct disorder score is 2.0 (2.3) for the sample of all twins and 1.3 (1.5) for the sample of identical twins. The intra class correlation for the conduct disorder score is 0.63 for identical twins and 0.30 for fraternal twins. The smaller amount of variation in the sample of identical twins makes the estimates more vulnerable to measurement error (Griliches, 1979) and may limit the opportunities for finding effects on human capital. There are more females than males in the data and this is quite normal for a sample of twins. It should also be noted that the number of pairs used in the estimation may be smaller due to missing values for human capital outcomes or one of the covariates.

Table 6.2 shows the distribution of the conduct disorder score for the sample of all twins (column 1) and for both categories of the other two definitions of conduct disorder. Nearly 39% of all twins reported negatively on all 21 statements of conduct disorder and nearly 73% have a conduct disorder score of less than three. This distribution is remarkably similar to the distribution of conduct disorder (antisocial/aggression) for Canada reported in Currie and Stabile (2007). As expected, the conduct disorder score for twins that reported ‘yes’ on the APA definition (column 3) or those who sought professional help (column 5) are much higher, on average more than 5 (4) points higher for the second (third) definition.

The frequencies of specific behaviours underlying the conduct disorder scores are shown in Table 6.3. Behaviours most frequently reported are: wagged school, stay out late, shoplifted, misbehaved at school, stole from home and family. Two of these (wagging school and staying out late) are also reported as most serious in Le et al. (2005). The largest differences between the categories of the second definition (columns 2 and 3) and the third

Table 6.1: Means table for sample of children with all conduct disorder measures non-missing

	All twins		Identical twins	
	Mean	Twin differences	Mean	Twin differences
Conduct disorder measures				
Conduct disorder score (1-21)	1.880	3596	1.674	1394
APA definition	0.131	956	0.112	330
Professional help	0.027	264	0.024	92
Positive human capital				
Grade repetition	0.157	792	0.167	250
Marks primary school (1-3)	2.337	2184	2.322	660
Marks high school (1-3)	2.268	2186	2.269	748
3+ school drop-out	0.018	58	0.016	22
Years of education	12.247	2872	12.284	1096
High school graduation	0.792	1102	0.79	197
Negative human capital				
3+ job quit	0.102	408	0.100	144
Lies	0.044	412	0.035	134
3+ Attacking others	0.192	658	0.191	252
Failed to pay debts	0.073	286	0.067	102
Arrested since 18	0.074	290	0.069	102
Jail	0.019	70	0.018	22
Covariates				
Education mother	10.337		10.320	
Education father	10.495		10.515	
Age in 1996	29.891		29.87	
Gender (male=1)	0.442		0.409	
Birth weight (grams)	2540		2430	
Sample size (N)	5322		2250	

Table 6.2: Distribution of conduct disorder score (% of twins with each score)

	All	APA definition		Professional help	
Score		No	Yes	No	Yes
0	38.67	44.81	0.00	39.78	0.00
1	20.37	23.61	0.00	20.72	8.28
2	13.62	15.79	0.00	13.78	8.28
3	8.89	8.23	10.12	8.78	12.41
4	5.62	3.71	17.30	5.60	6.21
5	3.56	1.75	14.81	3.45	7.59
6	3.15	1.09	17.01	2.85	13.10
7	1.85	0.55	10.56	1.63	8.97
8	1.22	0.20	8.21	1.10	5.52
9	1.17	0.20	7.77	0.93	9.66
10	0.81	0.07	5.87	0.60	8.28
11	0.38	0.00	2.93	0.35	1.38
12	0.40	0.00	3.08	0.23	6.21
13	0.15	0.00	1.17	0.12	1.38
14	0.02	0.00	0.15	0.02	0.00
15	0.09	0.00	0.73	0.04	2.07
17	0.04	0.00	0.29	0.02	0.69
Mean	1.88	1.18	6.46	1.75	6.30
N	5307	4579	682	5159	145

definition of conduct disorder (columns 4 and 5) are also found on these behaviours.

Table 6.3: Means of conduct disorder categories

	All	APA definition		Professional help	
Category		No	Yes	No	Yes
misbehaved in school	0.148	0.095	0.489	0.136	0.593
wagged school	0.336	0.260	0.806	0.327	0.728
suspended/expelled	0.091	0.062	0.277	0.085	0.335
stay out late	0.183	0.118	0.594	0.176	0.474
sneak out at night	0.078	0.032	0.374	0.071	0.358
run away overnight	0.064	0.028	0.241	0.056	0.366
lied, used false name	0.097	0.046	0.426	0.088	0.445
outsmarted, conned others	0.065	0.028	0.301	0.060	0.262
stole from home or family	0.150	0.102	0.460	0.143	0.457
shoplifted	0.179	0.122	0.552	0.171	0.486
forged signature	0.003	0.001	0.016	0.003	0.029
damaged property	0.090	0.048	0.364	0.086	0.283
started physical fights	0.069	0.037	0.277	0.065	0.237
used a weapon	0.027	0.015	0.108	0.025	0.127
physically injured someone	0.065	0.043	0.207	0.063	0.173
bullied others	0.028	0.013	0.128	0.025	0.133
mean to animals	0.032	0.020	0.105	0.031	0.064
lighted fires	0.141	0.105	0.370	0.138	0.272
broke into someoneŠs car/house	0.068	0.027	0.331	0.062	0.289
forcefully stole money or property	0.002	0.000	0.015	0.002	0.023
forced someone into sexual activity	0.001	0.000	0.010	0.001	0.006

A first exploration of the relationship between conduct disorder and human capital is shown in Figure 6.1. The figure shows Lowess plots of the association between the conduct disorder score and our measures of positive and negative human capital. Lowess regression carries out a locally weighted regression of y variable (human capital) on x variable (conduct disorder), where we used least-squares smoothing (default in Stata).

The association between conduct disorder and human capital seems remarkably consistent for all measures of human capital. An increase of conduct disorder is associated with a decrease in positive human capital and increase in negative human capital. For instance, higher levels of conduct disorder are associated with higher probabilities of grade repetition and drop out, and with lower marks in school (primary or high school), less years of education, and a lower probability of graduating from high school. In addition, higher levels of conduct disorder are associated with higher probabilities of lying, attacking others, being arrested since the age of 18 or having been incarcerated. In line with previous studies for the US and Canada (Currie and Stabile, 2006, 2007), we find that the association between conduct disorder and human capital seems quite linear. Even low scores of conduct disorder are associated with lower human capital accumulation.

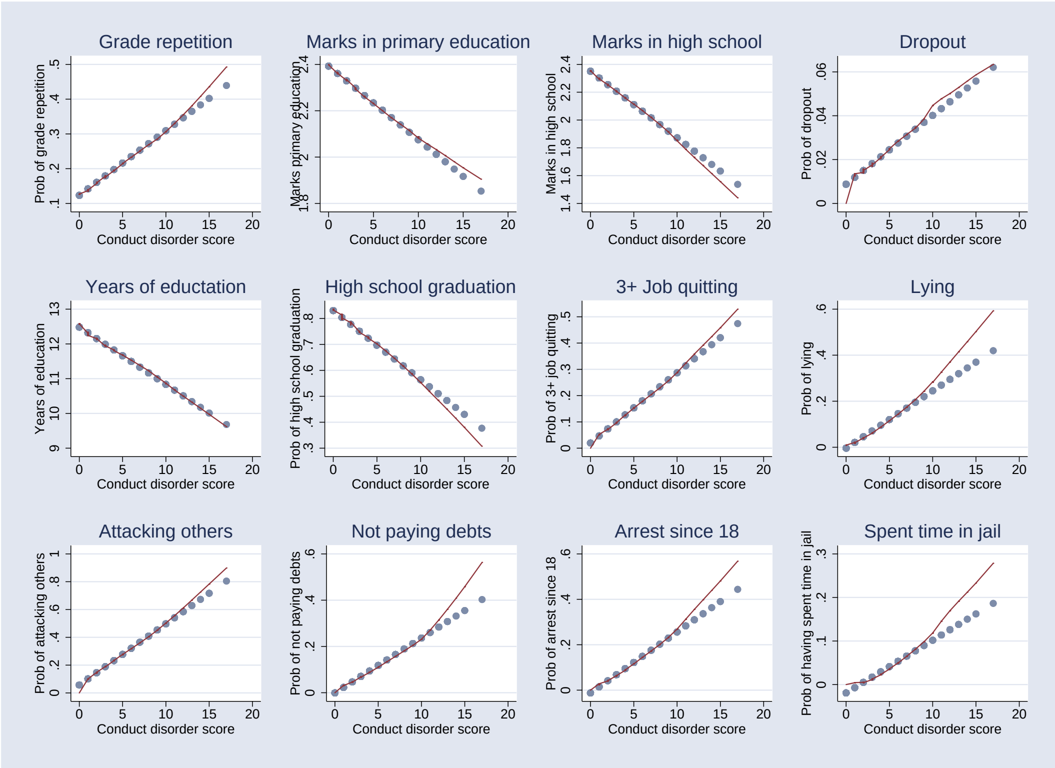

Figure 6.1: The associations between conduct disorder scores and human capital

This means that children with symptoms of conduct disorder are just as likely to suffer reduced human capital accumulation later in life as children who have been diagnosed with conduct disorder.

6.4 Methodology

We follow the same approach as in previous papers on the effects of childhood mental health problems on human capital. In order to introduce the methodology based on twin-differencing, we follow the exposition of Lundborg (2008). First, we estimate OLS models of the relationship between conduct disorder and positive and negative human capital outcomes. These estimates might be biased due to the unobserved heterogeneity between conduct disorder and human capital. To see this, consider an individual i that has positive/negative human capital HC_i determined by:

$$HC_i = \beta \; CD_i + \alpha F_i + u_i, \tag{6.1}$$

where CD_i denotes one of our three measures of conduct disorder, F_i denotes unobserved "family background", and u_i is an unobserved random component. We assume that "family background" denotes both unobserved genetic traits affecting human capital, as well as unobserved family circumstances. Further, let conduct disorder be determined by:

$$CD_i = \delta F_i + \xi_i, \tag{6.2}$$

where F_i denotes the same unobserved "family background" that affects human capital and ξ_i denotes conduct disorder-specific random term. Due to the unobserved genetic traits and family background characteristics affecting both conduct disorder and human capital measures, we get a standard result that OLS is upward biased:

$$plim(\beta_{OLS}) = \beta + \alpha \frac{\sigma_{F,CD}}{\sigma^2_{CD}}. \tag{6.3}$$

Therefore, we proceed with estimating within twin fixed effect model. Let HC_{1j} and HC_{2j} denote human capital outcome of the first and second twin in the jth twin pair. The unobserved component is again made up of two parts. The first part, μ_j, denotes unobserved factors that vary between twin pairs but not within pairs. The second part, ε_{1j} and ε_{2j}, denote unobserved factors specific to each twin. We then get:

$$HC_{1j} = \beta\ CD_{1j} + \mu_j + \varepsilon_{1j}, \tag{6.4}$$
$$HC_{2j} = \beta\ CD_{2j} + \mu_j + \varepsilon_{2j}, \tag{6.5}$$

Taking first differences between equations (6.4) and (6.5) gives:

$$HC_{1j} - HC_{2j} = \beta_{WTP}(CD_{1j} - CD_{2j}) + \varepsilon_{1j} - \varepsilon_{2j}, \tag{6.6}$$

where β_{WTP} is the within-twin-pair estimate of conduct disorder. Adding explanatory variables, we estimate the following equation using OLS and twin FE approaches:

$$HC_{ij} = \beta\ CD_{ij} + \gamma X_{ij} + \mu_j + \varepsilon_{ij} \tag{6.7}$$

where index i corresponds to the twin number, while index j corresponds to the twin family; X_{ij} denotes a vector of covariates, consisting of mother's and father's education, age and age squared, gender, and weight at birth. In the FE specification, all explanatory variables that are common to both twins in a given twin pair will be differenced out. Since twins have the same genetic "make-up", the influence of genes as well as the influence of common family circumstances will vanish. This means that an OLS estimate of equation (6.6) will no longer be biased due to unobserved characteristics specific to the twin-pair within one family.

We also use OLS and within-twin models for estimating the effect of the age at which conduct disorder behaviours occur on human capital accumulation. Our data contains information on the age of onset of the behaviours according to the APA definition. We created a variable 'years of conduct disorder' as the years before the age of 18 that these behaviours first occurred (18 − age of onset). Hence, twins coded as zero on the APA definition have a value of zero on this variable. Twins with age of onset of 17 have a value of one on this variable, twins with an age of onset of 16 have a value of 2 on this variable, and so on. For investigating the effect of age of onset of conduct disorder we estimated the following fixed effect model:

$$HC_{ij} = \beta(years\ cd_{ij}) + \delta(years\ cd_{ij})^2 + \gamma X_{ij} + \mu_j + \varepsilon_{ij} \tag{6.8}$$

There are several concern with estimates based on within-family models. First, the estimates might be biased by within-family heterogeneity. Most previous papers on mental health problems of children estimate within-family models using data of siblings. However, siblings may differ in genetic endowments. In addition the socioeconomic conditions facing siblings and the parental inputs received by siblings may differ if family circumstances change over time. As we use data on (identical) twins it seems less likely that our estimates will be biased by within family heterogeneity. The family circumstances for twins will probably be more equal than for siblings, which differ in age. More importantly, identical twins share exactly the same genes (have the same genetic "make-up"), whereas siblings on average only share half of their genetic endowments. By estimating fixed effect models for separate samples of identical twins we control for all differences in genetic endowments. In addition, we test the robustness of the estimates by excluding pairs of twins with very large differences in educational attainment. These large education difference might indicate that these twins are quantitatively different from the rest of the sample and introduce heterogeneity which will confound the effects we are looking for.

A second concern for our analysis is measurement error in conduct disorder. The FE estimate of β might still be biased due to the measurement error in conduct disorder. It is well-known from the literature that the within-family estimator exacerbates measurement error, which is likely to bias the estimates towards zero. The bias increases even more when differencing between identical twins (Griliches, 1979). Bound and Solon (1999) call this "double trouble." We address the measurement error problem by taking advantage of the rich set of indicators of conduct disorder in our data. First, we use three separate measures of conduct disorder. If these three measures yield comparable estimates, we can be more confident about the results. Second, we use two measures of conduct disorder in an instrumental variable approach. Several studies on the returns to schooling using samples of twins address measurement error in schooling by instrumenting with a second measure of schooling (Ashenfelter and Krueger, 1994; Miller et al., 1995). These studies typically show that instrumenting leads to higher estimates of the returns to schooling. We instrument the APA definition of conduct disorder (our second measure) with the conduct disorder score (our first measure of conduct disorder).

A third concern for employing within-family estimation is that there might be spill-over effects within pairs of twins. If severe conduct disorder behaviours of one twin also have negative effects on the human capital accumulation of the other twin, then using within twin estimation would underestimate the effect of conduct disorder. Spill-over effects seem quite likely here as conduct disorder is known as a 'disruptive behaviour disorder' because of its impact on children and their families, neighbours, and schools. We address this issue in the same way as Fletcher and Wolfe (2008). They included the mental problems of the other twin in a random effects regression controlling for mental problems. Although the estimates can be biased by omitted variables they might provide insights on spill-over effects.

6.5 Main estimation results

In this section we present the main estimation results. Section 6.5.1 shows the estimates of the effect of conduct disorder on human capital using three measures of conduct disorder. In Section 6.5.2 we investigate whether the age of onset of conduct disorder matters for human capital accumulation.

6.5.1 The effect of conduct disorder on human capital

We examine the effect of conduct disorder problems, as a measure of non-cognitive traits, on different human capital outcomes using three measures of conduct disorder. Tables 6.4 and 6.5 report the estimated effects of conduct disorder on 6 types of positive and negative human capital, respectively, using linear regression (probability) models. The top panel shows the results using the conduct disorder score, the middle panel shows the results for the APA definition of conduct disorder and the bottom panel shows the results using the professional help definition. In each panel we first report OLS estimates, next we report fixed effect estimates for the sample of all twins (FE All) and finally we report fixed effect estimates for the sample of identical twins (FE MZ). All regressions include birth weight, age, age squared, gender, and parents' education as controls. Obviously, some of these variables drop out from the fixed effect specifications. Each cell shows the results of a separate estimation.

All estimates in table 6.4 suggest that conduct disorder has a negative effect on human capital accumulation, increasing the probability of grade repetition and dropping out of school, and decreasing marks in school, years of education completed and the probability of graduating from high school. The largest effects are found when using the OLS estimation. The size of the estimates reduces when twin fixed effects are taken into account. For the sample of all twins all estimates, except for one, remain statistically significant. The standard errors increase when we restrict the sample to identical twins only. However, even for the sample of identical twins, where the variation in conduct disorder is much smaller than in the sample of all twins (see Table 6.1), several estimates suggest substantial effects of conduct disorder on human capital accumulation. Most remarkable are the effect on grade retention and high school completion. One additional point on the conduct disorder score increase the probability of grade retention with 1.1 percentages point (0.9 for the sample of all twins). This estimate is remarkably close to the within sibling estimate of 0.8 percentage points both for the US and Canada reported by Currie and Stabile (2007). Conduct disorder as classified by the APA definition increases the probability of grade retention with 9 percentage points; for the third definition of conduct disorder, the estimated effect is 11 percentage points. The estimates of the effects on high school completion are also large although not always statistically significant in the sample of identical twins. The estimates for the third definition suggest that conduct disorder decreases the probability of high school graduation with 13 percentage points. The size of these estimates seems large when compared with other covariates. For instance, in the OLS regression we find that one year of father's education is associated with a decrease of the probability of grade retention with 0.5% and an increase of the probability of high school graduation with 2.2%.

The pattern of findings in Table 6.5 is quite similar and, considering the statistical significance of the estimates, even more compelling. We find large and statistically significant effects of conduct disorder on negative human capital, even in the sample of identical twins. The estimates of the effects on the probability of being arrested since the age of 18 and on the probability of spending time in jail seems quite large and robust. In addition,

Table 6.4: Estimates of the effect of conduct disorder on positive human capital

	Grade retention	Marks primary school	Marks high school	3+ Drop-out	Education years	High school
Conduct Disorder Score						
OLS	0.017	-0.029	-0.049	0.005	-0.160	-0.026
	(0.003)***	(0.004)***	(0.004)***	(0.002)**	(0.013)***	(0.003)***
N	5286	5276	5270	1732	5288	5288
FE All	0.009	-0.024	-0.038	0.005	-0.082	-0.013
	(0.003)***	(0.005)***	(0.005)***	(0.002)**	(0.018)***	(0.003)***
N	5286	5276	5270	1732	5288	5288
FE MZ	0.011	-0.014	-0.017	0.009	-0.037	-0.010
	(0.005)**	(0.008)*	(0.009)**	(0.004)**	(0.032)	(0.006)
N	2238	2234	2238	720	2236	2236
APA Definition						
OLS	0.096	-0.171	-0.267	0.029	-0.818	-0.136
	(0.019)***	(0.026)***	(0.026)***	(0.011)**	(0.095)***	(0.019)***
N	5224	5214	5210	1700	5226	5226
FE All	0.057	-0.100	-0.180	0.031	-0.335	-0.054
	(0.018)***	(0.032)***	(0.031)***	(0.011)***	(0.111)***	(0.020)***
N	5224	5214	5210	1700	5226	5226
FE MZ	0.091	-0.031	-0.025	0.028	-0.095	-0.036
	(0.026)***	(0.043)	(0.045)	(0.017)	(0.169)	(0.032)
N	2220	2216	2220	714	2218	2218
Professional help						
OLS	0.201	-0.194	-0.308	0.045	-1.135	-0.218
	(0.039)***	(0.054)***	(0.049)***	(0.026)*	(0.187)***	(0.040)***
N	5370	5360	5354	1794	5372	5372
FE All	0.105	-0.102	-0.157	0.014	-0.661	-0.126
	(0.033)***	(0.060)*	(0.059)***	(0.021)	(0.208)***	(0.038)***
N	5370	5360	5354	1794	5372	5372
FE MZ	0.109	-0.067	-0.088	-0.034	-0.480	-0.130
	(0.048)**	(0.082)	(0.085)	(0.032)	(0.317)	(0.061)**
N	2274	2270	2274	744	2272	2272

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

Table 6.5: Estimates of the effect of conduct disorder on negative human capital

	3+ Job quits	Lies	Attacking others	Debts	Arrested since 18	Jail
Conduct Disorder Score						
OLS	0.025	0.027	0.041	0.026	0.026	0.014
	(0.003)***	(0.002)***	(0.004)***	(0.003)***	(0.003)***	(0.003)***
N	2182	5288	2180	2182	2178	2176
FE All	0.022	0.028	0.037	0.027	0.019	0.011
	(0.004)***	(0.002)***	(0.006)***	(0.004)***	(0.004)***	(0.002)***
N	2182	5288	2180	2182	2178	2176
FE MZ	0.009	0.021	0.029	0.031	0.011	0.007
	(0.008)	(0.004)***	(0.011)***	(0.007)***	(0.007)	(0.003)**
N	888	2234	888	888	888	886
APA Definition						
OLS	0.103	0.162	0.179	0.088	0.124	0.048
	(0.018)***	(0.016)***	(0.022)***	(0.016)***	(0.017)***	(0.011)***
N	2142	5226	2140	2142	2138	2136
FE All	0.061	0.128	0.146	0.051	0.076	0.02
	(0.023)***	(0.013)***	(0.029)***	(0.019)***	(0.019)***	(0.010)**
N	2142	5226	2140	2142	2138	2136
FE MZ	0.038	0.091	0.162	0.037	0.067	0.022
	(0.034)	(0.019)***	(0.044)***	(0.028)	(0.028)**	(0.013)*
N	876	2216	876	876	876	874
Professional help						
OLS	0.089	0.112	0.153	0.15	0.125	0.137
	(0.038)**	(0.030)***	(0.045)***	(0.040)***	(0.038)***	(0.035)***
N	2248	5370	2246	2248	2244	2242
FE All	0.022	0.053	0.022	0.03	0.061	0.101
	(0.042)	(0.024)**	(0.054)	(0.035)	(0.036)*	(0.017)***
N	2248	5370	2246	2248	2244	2242
FE MZ	-0.002	0.065	0.001	0.024	0.157	0.081
	(0.064)	(0.036)*	(0.085)	(0.054)	(0.054)***	(0.025)***
N	916	2268	916	916	916	914

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

the estimates of the effects of conduct disorder on the probability of physically attacking others are very large for the first two measures of conduct disorder. For the third measure the effects are only found in the OLS estimates. It should be noted that the sample size is much smaller for five measures of negative human capital. This is caused by the routing of the questionnaire and may bias the results. In the next section we will investigate the sensitivity of the results with respect to the smaller sample size. In general, we find the largest effects of conduct disorder for the third measure (Professional help). It is likely that this measure identifies individuals with severe behaviour problems which hinder their development. The size of the estimates for the first and the second measure seems roughly comparable if we consider that twins with conduct disorder according to the APA definition score approximately 5 points higher on the conduct disorder score.

Non-linearity

We also investigated whether the effects are non-linear by looking at the top percentiles of conduct disorder score. Tables 6.6 and 6.7 show the estimation results of a dummy for having a conduct disorder score of more than 5 or at least 10. This corresponds to the 90th and 99th percentile. It should be noted that focusing on these top percentiles reduces the variation in conduct disorder that can be used, especially in the sample of identical twins. For most outcomes the estimates suggest that the effects of conduct disorder are quite linear. For instance, for the sample of all twins the within estimates using the dummy for at least 10 conduct disorder behaviours (above the 98th percentile) seem quite comparable with most estimates in Table 6.5. Hence, we find no clear evidence for non-linear effects of conduct disorder.

Effects of four subscales of conduct disorder

The APA criteria (see Section 6.A.2 in the Appendix 6.A) suggest that conduct disorder has four underlying components: aggression to people and animals, destruction of property, deceitfulness or theft, and serious violations of rules. We investigated the effects of these components by decomposing the conduct disorder score into these four groups of the APA criteria. The estimation results are shown in Tables 6.8 and 6.9. The results in Table 6.8 suggest that 'serious violations of rules' is the most important component of conduct disorder for the accumulation of positive human capital. This not very surprising if we consider that this subscale includes categories like 'misbehaved', 'wagged school', and 'suspended/expelled from school'. For the other subscales, the estimates do not show a clear pattern. The estimates of the effect on negative human capital (Table 6.9) suggest that all subscales are important. The size of the effects of 'aggression' and 'destruction' on 'attacking others' and 'failing to pay debts' is remarkable. 'Destruction', 'deceitfulness/theft' and 'violation of rules' seem the most important factors for the probability of arrest or jail.

Summary of main estimation results

Using a rich set of indicators of conduct disorder we find large effects of conduct disorder on positive and negative human capital. Even for the sample of identical twins we

Table 6.6: Estimates of the effect of more than 5 conduct disorder behaviours (above 90^{th} percentile)

CD>5	Grade retention	Marks primary	Marks high school	3+ dropout	Education years	High school
OLS	0.103	-0.169	-0.303	0.020	-0.897	-0.158
	(0.022)***	(0.033)***	(0.032)***	-0.013	(0.111)***	(0.024)***
N	5286	5276	5270	1732	5288	5288
FE All	0.032	-0.106	-0.207	0.018	-0.372	-0.073
	(0.022)	(0.039)***	(0.039)***	(0.014)	(0.136)***	(0.025)***
N	5286	5276	5270	1732	5288	5288
FE MZ	0.069	-0.069	-0.103	0.040	-0.207	-0.069
	(0.035)*	(0.059)	(0.062)*	(0.025)	(0.232)	(0.044)
N	2238	2234	2238	720	2236	2236
	3+ Job quits	Lies	Attacking others	Debts	Arrested since 18	Jail
OLS	0.132	0.192	0.198	0.131	0.137	0.075
	(0.022)***	(0.019)***	(0.026)***	(0.020)***	(0.021)***	(0.015)***
N	2182	5288	2180	2182	2178	2176
FE All	0.099	0.185	0.139	0.131	0.085	0.039
	(0.028)***	(0.015)***	(0.035)***	(0.023)***	(0.023)***	(0.011)***
N	2182	5288	2180	2182	2178	2176
FE MZ	0.001	0.128	0.041	0.112	0.014	0.028
	(0.047)	(0.026)***	(0.062)	(0.039)***	(0.039)	(0.018)
N	888	2234	888	888	888	886

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

Table 6.7: Estimates of the effect of at least 10 conduct disorder behaviours (above 98$^{\text{th}}$ percentile)

CD ≥10	Grade retention	Marks primary	Marks high school	3+ dropout	Education years	High school
OLS	0.236	-0.227	-0.474	0.039	-1.335	-0.267
	(0.050)***	(0.069)***	(0.071)***	-0.032	(0.201)***	(0.050)***
N	5286	5276	5270	1732	5288	5288
FE All	0.131	-0.07	-0.201	0.022	-0.617	-0.148
	(0.041)***	(0.075)	(0.074)***	(0.027)	(0.260)**	(0.048)***
N	5286	5276	5270	1732	5288	5288
FE MZ	0.169	-0.061	-0.060	0.092	0.134	0.003
	(0.078)**	(0.131)	(0.136)	(0.052)*	(0.510)	(0.097)
N	2238	2234	2238	720	2236	2236

	3+ Job quits	Lies	Attacking others	Debts	Arrested since 18	Jail
OLS	0.171	0.343	0.384	0.249	0.257	0.157
	(0.051)***	(0.050)***	(0.055)***	(0.051)***	(0.052)***	(0.042)***
N	2182	5288	2180	2182	2178	2176
FE All	0.068	0.270	0.285	0.185	0.163	0.131
	(0.051)	(0.029)***	(0.065)***	(0.043)***	(0.043)***	(0.021)***
N	2182	5288	2180	2182	2178	2176
FE MZ	-0.123	0.055	0.237	0.232	-0.002	0.126
	(0.097)	(0.056)	(0.127)*	(0.081)***	(0.080)	(0.037)***
N	888	2234	888	888	888	886

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

Table 6.8: The effect of four subscales of conduct disorder on positive human capital

Subscales of conduct disorder	Grade retention	Marks primary school	Marks high school	3+ Drop-out	Education years	High school
OLS						
Aggression	0.005	0.008	0.014	0.001	0.052	-0.019
	(0.012)	(0.017)	(0.017)	(0.004)	(0.061)	(0.012)*
Destruction	0.012	0.045	0.046	0.002	0.052	0.019
	(0.013)	(0.019)**	(0.018)**	(0.005)	(0.070)	(0.013)
Deceit/theft	0.009	-0.020	-0.016	0.005	0.007	-0.005
	(0.007)	(0.011)*	(0.010)	(0.003)	-0.039	(0.007)
Violation of rules	0.029	-0.065	-0.119	0.006	-0.409	-0.056
	(0.006)***	(0.008)***	(0.008)***	(0.002)**	(0.029)***	(0.006)***
N	5286	5276	5270	4642	5288	5288
FE All						
Aggression	-0.011	-0.006	0.028	-0.007	0.035	-0.009
	(0.011)	(0.020)	(0.019)	(0.004)*	(0.068)	(0.013)
Destruction	0.009	0.015	0.019	0.006	0.005	0.011
	(0.013)	(0.024)	(0.023)	(0.005)	(0.083)	(0.015)
Deceit/theft	-0.001	-0.022	-0.035	0.006	-0.072	-0.005
	(0.007)	(0.013)	(0.013)***	(0.003)**	(0.046)	(0.008)
Violation of rules	0.024	-0.044	-0.081	0.006	-0.159	-0.029
	(0.006)***	(0.011)***	(0.011)***	(0.002)**	(0.039)***	(0.007)***
N	5286	5276	5270	4642	5288	5288
FE MZ						
Aggression	0.013	-0.020	0.001	-0.003	-0.020	0.003
	(0.016)	(0.027)	(0.028)	(0.007)	(0.106)	(0.020)
Destruction	-0.004	-0.006	-0.046	0.017	-0.070	-0.002
	(0.019)	(0.032)	(0.033)	(0.008)**	(0.125)	(0.024)
Deceit/theft	0.010	0.013	0.038	0.011	-0.037	-0.024
	(0.011)	(0.019)	(0.019)*	(0.005)**	(0.073)	(0.014)*
Violation of rules	0.016	-0.035	-0.061	0.001	-0.034	-0.006
	(0.009)*	(0.016)**	(0.016)***	(0.004)	(0.062)	(0.012)
N	2238	2234	2238	1992	2236	2236

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

Table 6.9: The effect of four subscales of conduct disorder on negative human capital

Subsacles of conduct disorder	3+ Job quits	Lies	Attacking others	Debts	Arrested since 18	Jail
OLS						
Aggression	0.004	0.036	0.116	0.028	0.016	0.015
	(0.006)	(0.012)***	(0.015)***	(0.011)**	(0.012)	(0.008)*
Destruction	-0.004	0.014	0.040	0.019	0.025	0.012
	(0.006)	(0.013)	(0.016)**	(0.011)*	(0.013)**	(0.007)*
Deceit/theft	0.008	0.025	0.031	0.024	0.030	0.012
	(0.005)*	(0.007)***	(0.009)***	(0.007)***	(0.006)***	(0.004)***
Violation of rules	0.005	0.024	0.020	0.030	0.027	0.015
	(0.004)	(0.006)***	(0.007)***	(0.005)***	(0.005)***	(0.003)***
N	1732	2182	2180	2182	2178	2176
FE All						
Aggression	-0.008	0.038	0.100	0.032	0.007	0.004
	(0.007)	(0.014)***	(0.018)***	(0.012)***	(0.012)	(0.006)
Destruction	0.004	0.018	0.062	0.021	0.016	-0.006
	(0.009)	(0.018)	(0.022)***	(0.0150	(0.015)	(0.007)
Deceit/theft	0.010	0.028	0.032	0.035	0.020	0.009
	(0.005)**	(0.010)***	(0.012)***	(0.008)***	(0.008)**	(0.004)**
Violation of rules	0.006	0.008	0.004	0.019	0.024	0.019
	(0.004)	(0.009)	(0.011)	(0.007)**	(0.008)***	(0.004)***
N	1732	2182	2180	2182	2178	2176
FE MZ						
Aggression	-0.002	0.005	0.080	0.051	-0.000	-0.005
	(0.011)	(0.022)	(0.029)***	(0.018)***	(0.018)	(0.008)
Destruction	0.023	-0.002	0.116	0.042	0.030	-0.004
	(0.014)*	(0.028)	(0.037)***	(0.023)*	(0.023)	(0.011)
Deceit/theft	0.017	-0.006	0.013	0.023	0.010	0.007
	(0.008)**	(0.016)	(0.021)	(0.013)*	(0.013)	(0.006)
Violation of rules	0.004	0.028	-0.009	0.024	0.012	0.017
	(0.008)	(0.015)*	(0.019)	(0.012)**	(0.012)	(0.006)***
N	720	888	888	888	888	886

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

find substantial effects. These findings corroborate findings from previous studies based on within-sibling estimation and using smaller sets of indicators of conduct disorder. Although the within-twin estimates suggest substantial deleterious effects of conduct disorder, these estimates are considerably smaller than the OLS-estimates. This difference between the OLS estimates and the within-twin estimates might result from measurement error in conduct disorder or spill-over effects within pairs of twins. We will address these issues in Section 6.6.

6.5.2 The effect of the timing of conduct disorder behaviours

Several recent studies suggest that the timing of intervention programmes for disadvantaged children is important. For instance, Carneiro and Heckman (2003) show that socio-economic differences in cognitive and non-cognitive abilities appear early in life and widen over the life cycle of the child. Currie and Stabile (2007) find that early mental health problems have large significant effects on cognitive test scores even controlling for later mental health problems. Loeber (1982) suggests that the earlier a child began displaying antisocial behaviour, the more likely he or she was to persist in such behaviour. Farmer (1995), on the other hand, does not support the hypothesis that earlier onset is associated with more deleterious outcomes. She finds that later onset, rather than earlier, is associated with earlier school leaving and lower level qualifications. In this section we analyse the impact of the timing of conduct disorder behaviours.

Our data contains information on the age of onset of the behaviours according to the APA definition. We created a variable 'onset of conduct disorder' as the years before the age of 18 that these behaviours first occurred (18 minus age of onset). The age of onset of conduct disorder appears to be strongly associated with the number of conduct disorder behaviours manifested before the age of 18 (see Table 6.10). More years of conduct disorder is associated with a higher conduct disorder score. We estimated the effect of the timing of conduct disorder on human capital by including 'years of conduct disorder' and 'years of conduct disorder squared' in our main models (equation (6.8) in Section 6.4). Tables 6.11 and 6.12 show the estimation results.

Table 6.10: The effect of the timing of conduct disorder on the conduct disorder score

	OLS	FE	
		All	Identical
Years of CD	1.883	1.397	1.035
	(0.080)***	(0.050)***	(0.069)***
Years of CD^2	-0.134	-0.102	-0.073
	(0.012)***	(0.007)***	(0.010)***
N	5198	5198	2208

The estimates suggest that the timing of conduct disorder behaviours is very important for human capital accumulation. All the estimates corroborate the previous findings and indicate that the effect of conduct disorder on human capital is more deleterious if

Table 6.11: The effect of the timing of conduct disorder on positive human capital

APA Definition	Grade retention	Marks primary school	Marks high school	3+ Drop-out	Education years	High school
OLS						
Years of CD	0.040	-0.060	-0.110	0.013	-0.333	-0.058
	(0.009)***	(0.014)***	(0.013)***	(0.006)**	(0.043)***	(0.009)***
Years of CD^2	-0.003	0.004	0.009	-0.001	0.026	0.004
	(0.001)***	(0.002)**	(0.002)***	(0.001)	(0.005)***	(0.001)***
N	5222	5212	5208	1698	5224	5224
FE All						
Years of CD	0.024	-0.041	-0.083	0.023	-0.153	-0.036
	(0.009)***	(0.016)**	(0.016)***	(0.006)***	(0.056)***	(0.010)***
Years of CD^2	-0.002	0.004	0.008	-0.003	0.013	0.004
	(0.001)*	(0.002)*	(0.002)***	(0.001)***	(0.007)*	(0.001)***
N	5222	5212	5208	1698	5224	5224
FE MZ						
Years of CD	0.044	0.015	0.000	0.026	-0.009	-0.033
	(0.013)***	(0.023)	(0.024)	(0.010)***	(0.088)	(0.017)*
Years of CD^2	-0.004	-0.004	-0.001	-0.003	-0.001	0.005
	(0.002)**	(0.003)	(0.003)	(0.001)**	(0.012)	(0.002)*
N	2220	2216	2220	714	2218	2218

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

Table 6.12: The effect of the timing of conduct disorder on negative human capital

APA Definition	3+ Job quits	Lies	Attacking others	Debts	Arrested since 18	Jail
OLS						
Years of CD	0.027	0.057	0.061	0.033	0.042	0.012
	(0.010)***	(0.009)***	(0.011)***	(0.009)***	(0.009)***	(0.007)*
Years of CD^2	-0.000	-0.004	-0.004	-0.002	-0.003	0.000
	(0.001)	(0.001)***	(0.002)**	(0.001)*	(0.001)**	(0.001)
N	2140	5224	2138	2140	2136	2134
FE All						
Years of CD	0.005	0.049	0.060	0.012	0.035	0.011
	(0.011)	(0.006)***	(0.014)***	(0.009)	(0.009)***	(0.005)**
Years of CD^2	0.002	-0.004	-0.005	-0.000	-0.003	-0.001
	(0.001)	(0.001)***	(0.002)**	(0.001)	(0.001)***	(0.001)**
N	2140	5224	2138	2140	2136	2134
FE MZ						
Years of CD	0.006	0.042	0.056	0.019	0.032	0.004
	(0.017)	(0.010)***	(0.023)**	(0.014)	(0.014)**	(0.007)
Years of CD^2	0.001	-0.003	-0.003	-0.002	-0.003	0.000
	(0.002)	(0.001)**	(0.003)	(0.002)	(0.002)*	(0.001)
N	876	2216	876	876	876	874

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

these behaviour occur earlier. Even in the sample of identical twins we find statistically significant effects of the timing of conduct disorder on grade retention, drop out, high school graduation, attacking other people and the probability of being arrested since the age of 18. The impact of the timing differs between the outcomes. For instance, the quadratic specification suggests that conduct disorder is most deleterious for high school completion when manifested between the ages from 13 to 16. In addition, the effect of conduct disorder on the probability of arrest since the age of 18 is largest when conduct disorder manifests itself between the ages from 10 to 16.

We also investigated whether there is a relationship between the socioeconomic background of the twins and the effect of the timing of conduct disorder on human capital. As wealthier families have more resources to mitigate conduct disorder problems we might expect smaller effects for twins from these families. We separated our sample in families with a father that completed high school and families with a dad that did not complete high school. However, estimations which include an interaction between socioeconomic background and years of conduct disorder show little effect of socioeconomic background. We only find that conduct disorder has a larger effect on the probability of physically attacking others in families in which the dad did not complete high school than in other families. Currie and Stabile (2007) also find little evidence that parents' income mitigates negative effects of mental health problems of children.

6.5.3 Summary of the timing of conduct disorder

We conclude that the timing of conduct disorder is very important for human capital accumulation. In line with previous studies, this suggests that programmes that succeed in changing conduct disorder behaviours at an early age might yield large returns, both for individuals and society at large.

6.6 Robustness checks

In this section we investigate the robustness of the findings by addressing four issues. First, we test the sensitivity of the results by imputing missing values on human capital outcomes which are due to the routing of the questionnaire. Second, we exclude pairs of twins with large differences in educational attainment. These large differences might indicate that there are other major differences between these twins. Third, we address the issue of measurement error in conduct disorder which is likely to bias the estimates downward. Fourth, we investigate whether spill-over effects of conduct disorder within pairs of twins play a role. In the robustness checks we mainly use the APA definition of conduct disorder because of the precise match of this measure with the psychiatric definition.

6.6.1 Missing values due to the routing of the questionnaire

In the previous section we noted that the number of observations is much smaller in the estimations of the effect on dropout and five measures of negative human capital. Due to the routing of the questionnaire twins with a conduct disorder score of zero did not answer these questions. This may bias the estimates downward as it seems likely that twins with no childhood conduct disorder behaviour will on average accumulate more positive and less negative human capital than twins with problem behaviour early in life. We checked the sensitivity of the results by imputing zeros for twins with missing values on these outcomes and a conduct disorder score of zero. Table 6.13 shows the estimation results for the APA definition of conduct disorder. After the imputation of these missing values all estimates become statistically significant. We also observe that the size of most estimates increases. This suggest that due to the routing of the questionnaire we underestimate the deleterious effects of conduct disorder on negative human capital. Hence, the size of the estimates in the previous section may be considered as lower bound of the true effects of conduct disorder on these outcomes.

Table 6.13: Estimates of the effect of conduct disorder (APA definition) on positive and negative human capital after imputation of missing values due to the routing of the questionnaire

APA Definition	Drop-out	Job quit	Attack others	Debts	Arrested	Jail
OLS	0.029	0.154	0.259	0.109	0.147	0.051
	(0.009)***	(0.016)***	(0.019)***	(0.014)***	(0.015)***	(0.010)***
N	4586	5184	5184	5184	5180	5178
FE All	0.025	0.108	0.201	0.066	0.100	0.029
	(0.007)***	(0.015)***	(0.019)***	(0.012)***	(0.012)***	(0.006)***
N	4586	5184	5184	5184	5180	5178
FE MZ	0.023	0.072	0.200	0.060	0.073	0.025
	(0.011)**	(0.023)***	(0.030)***	(0.019)***	(0.018)***	(0.008)***
N	1980	2198	2200	2200	2198	2196

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

6.6.2 Excluding pairs of twins with large differences in education

Second, we excluded pairs of twins with large differences in educational attainment. The large differences in education might indicate that these twins are quantitatively different from the rest of the sample and introduce heterogeneity which will confound the effects we are looking for. Table 6.14 shows estimates of the effect of the APA definition of conduct disorder on positive and negative human capital after excluding pairs of twins that differ by more than 5.5 years of education. The estimates in Table 6.14 are quite similar to the estimates in the previous section. Hence, the results are robust for exclusion of these pairs of twins.

Table 6.14: Estimates of the effect of conduct disorder (APA Definition) on positive and negative human capital after excluding pairs with extreme education differences

APA Definition	Grade retention	Marks primary school	Marks high school	3+ Drop-out	Education years	High school
OLS	0.095	-0.176	-0.264	0.030	-0.810	-0.137
	(0.019)***	(0.027)***	(0.027)***	(0.012)**	(0.095)***	(0.019)***
N	5076	5070	5066	1642	5078	5078
FE All	0.051	-0.104	-0.166	0.031	-0.281	-0.048
	(0.018)***	(0.032)***	(0.031)***	(0.012)***	(0.101)***	(0.019)**
N	5076	5070	5066	1642	5078	5078
FE MZ	0.086	-0.032	-0.019	0.028	-0.096	-0.037
	(0.026)***	(0.043)	(0.045)	(0.017)	(0.159)	(0.031)
N	1094	1092	1094	350	1093	1093

APA Definition	3+ Job quits	Lies	Attacking others	Debts	Arrested since 18	Jail
OLS	0.105	0.159	0.178	0.088	0.121	0.046
	(0.018)***	(0.016)***	(0.022)***	(0.016)***	(0.017)***	(0.011)***
N	2070	5078	2068	2070	2066	2064
FE All	0.061	0.121	0.148	0.048	0.076	0.020
	(0.023)***	(0.013)***	(0.029)***	(0.019)**	(0.020)***	(0.010)**
N	2070	5078	2068	2070	2066	2064
FE MZ	0.038	0.086	0.164	0.030	0.068	0.022
	(0.034)	(0.019)***	(0.045)***	(0.029)	(0.028)**	(0.013)*
N	860	2184	860	860	860	858

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

6.6.3 Measurement error in conduct disorder

As a third robustness check we investigate the impact of measurement error. It is well-known that the within-family estimator exacerbates measurement error, which is likely to bias the estimates towards zero. We address this issue by using instrumental variable estimation. Several studies on the returns to schooling using samples of twins address measurement error in schooling by instrumenting with a second measure of schooling (Ashenfelter and Krueger, 1994; Miller et al., 1995). These studies obtained two measures of the schooling of one twin by asking the twins to report both on their own schooling as well as on the schooling of their co-twin. The second measure of schooling can be used as an instrument to correct for measurement error. These studies typically find that instrumenting leads to higher estimates of the returns to schooling. Unfortunately, our data do not include such a second measure of conduct disorder. However, we have information on 21 statements about conduct disorder behaviours and information whether at least three of these behaviours occurred within 12 months. We use the conduct disorder score (our first measure of conduct disorder) as an instrument for the APA definition of conduct disorder (our second measure). The correlation between these two measures of conduct disorder, which indicates the reliability ratio, is 0.73. Table 6.15 shows the instrumental variables (IV) estimates of the APA definition of conduct disorder after instrumenting with the conduct disorder score.

The size of the estimated effects of the APA definition of conduct disorder on positive and negative human capital increases after instrumenting for measurement error. Now we even observe large negative effects of conduct disorder in the sample of identical twins that are statistically significant. For instance, conduct disorder decreases the probability of high school graduation with 11.5 percentage points and increases the probability of arrest since 18 with 9 percentage points. These findings suggests that measurement error is important and has a downward bias on the estimates.

6.6.4 Spill-over effects within pairs of twins

In his book on Happiness, Richard Layard (2005) writes that it is very likely that if one identical twin has a certain positive (negative) trait, the other twin copies this behaviour, simply because identical twins have the same genes. He shows that percent of identical twins who are both juvenile delinquents is 91 percent, compared to 73 percent of non-identical twins. Also, percent of identical twins who are both criminal convicts as adults is 52 percent compared to 23 percent of non-identical twins. Similarly to Layard (2005), we look at how many identical and non-identical twins answer the following question "If one twin has a problem, does the other twin have it also?" The results are presented in Table 6.16. It is evident that for the identical twins, if one twin has repeated a grade, dropped out of school, attacked other people or has been arrested, it is more likely that the other twin copies this behaviour. Hence, the copying of negative traits or spill-over effects are important in our sample of twins and it is more likely to happen among the identical twins who share identical genetic make-up, than among the non-identical twins.

As a fourth robustness check we investigate the spill-over effects in the sample of both identical and non-identical twins. Our within-twin estimates will be biased downward if

Table 6.15: IV-estimates of the effect of conduct disorder on positive and negative human capital

APA Definition	Grade retention	Marks primary school	Marks high school	3+ Drop-out	Education years	High school
OLS	0.170	-0.294	-0.493	0.038	-1.585	-0.259
	(0.027)***	(0.039)***	(0.039)***	(0.017)**	(0.137)***	(0.028)***
N	5209	5199	5195	1688	5211	5211
FE All	0.089	-0.263	-0.383	0.041	-0.839	-0.143
	(0.029)***	(0.054)***	(0.053)***	(0.020)**	(0.187)***	(0.034)***
N	5209	5199	5195	1688	5211	5211
FE MZ	0.116	-0.148	-0.160	0.074	-0.395	-0.115
	(0.048)**	(0.082)*	(0.085)*	(0.033)**	-0.319	(0.061)*
N	2214	2210	2214	709	2212	2212

APA Definition	3+ Job quits	Lies	Attacking others	Debts	Arrested since 18	Jail
OLS	0.211	0.268	0.353	0.223	0.223	0.120
	(0.030)***	(0.022)***	(0.033)***	(0.028)***	(0.028)***	(0.022)***
N	2128	5211	2126	2128	2124	2122
FE All	0.190	0.290	0.327	0.233	0.168	0.095
	(0.040)***	(0.022)***	(0.051)***	(0.035)***	(0.034)***	(0.017)***
N	2128	5211	2126	2128	2124	2122
FE MZ	0.068	0.220	0.245	0.248	0.088	0.065
	(0.064)	(0.035)***	(0.084)***	(0.057)***	(0.052)*	(0.024)***
N	870	2210	870	870	870	868

Note: All = identical + non-identical twins; MZ = identical twins; The APA definition of conduct disorder has been instrumented with the conduct disorder score; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

Table 6.16: If one twin has a problem, does the other twin have it also? (Percent)

	Identical twins	Non-identical twins
Grade retention	11.17	6.46
School drop-out (3+)	0.28	0.20
Less than high school	12.35	9.97
Job quits (3+)	2.74	1.26
Lies	0.54	0.60
Attacking others	7.31	4.43
Debts	1.60	1.42
Arrested since 18	2.51	0.63
Jail	0.92	0.32

conduct disorder behaviours of one twin also have negative effects on the other twin. We investigated this issue by including the effect of the conduct disorder of the other twin in a random effects regression controlling for own conduct disorder. This approach has also been applied by Fletcher and Wolfe (2008). The estimates are shown in Table 6.17.

The estimates suggest that the sibling's conduct disorder is important for human capital accumulation. Nearly all the estimates of the effect of the conduct disorder of the other twin are statistically significant and substantially increase the deleterious effects of conduct disorder on human capital. For some measures of human capital, years of education, high school graduation, spent time in jail, we find that the estimated effect of the other twins conduct disorder is approximately half the size of the estimated effect of the own conduct disorder. Although these estimates might be biased by unobserved differences they indicate that spill-over effects within pairs of twins might be important.

6.6.5 Summary of robustness checks

This section investigated the robustness of the estimates in Section 6.5. We find that imputations of missing values due to the routing of the questionnaire increase the size of the estimates of conduct disorder on negative human capital. Excluding pairs of twins with large differences in educational attainment does not change the main findings. In addition, measurement error in conduct disorder might be important. The size of the estimates increases after instrumenting for measurement error. Finally, we find evidence for spill-over effects of conduct disorder within pairs of twins. If anything, these robustness checks suggest that the estimated effects of conduct disorder on human capital in the previous section might be lower bounds of the true effects.

6.7 Conclusions

In this chapter we examined the relationship between conduct disorder problems before the age of 18 and human capital accumulation. We used three measures of conduct

Table 6.17: Random effect estimates of own and sibling's conduct disorder on human capital

APA Definition	Grade retention	Marks primary school	Marks high school	3+ Drop-out	Education years	High school
Own CD	0.091	-0.159	-0.253	0.029	-0.735	-0.122
	(0.015)***	(0.024)***	(0.023)***	(0.008)***	(0.092)***	(0.016)***
Sibling's CD	0.037	-0.052	-0.071	-0.001	-0.378	-0.064
	(0.014)**	(0.023)**	(0.023)***	-0.008	(0.091)***	(0.016)***
N	5224	5214	5210	1700	5226	5226

APA Definition	3+ Job quits	Lies	Attacking others	Debts	Arrested since 18	Jail
Own CD	0.099	0.156	0.175	0.084	0.118	0.045
	(0.016)***	(0.008)***	(0.020)***	(0.013)***	(0.013)***	(0.007)***
Sibling's CD	0.029	0.034	0.029	0.032	0.044	0.025
	(0.015)*	(0.008)***	-0.019	(0.013)**	(0.013)***	(0.007)***
N	2142	5226	2140	2142	2138	2136

Note: All = identical + non-identical twins; MZ = identical twins; Standard errors in parentheses; Significant at 1% (***), 5% (**), and 10% (*)-level.

disorder: a ‘conduct disorder score’ based on 21 statements about behaviour problems, the APA definition (‘did you do at least three of these behaviours within 12 months?’) and an indicator of ‘professional help’ because of conduct disorder. We estimated the effect of conduct disorder on six measures of positive human capital and six measures of negative human capital. The estimates suggest that conduct disorder has a strong negative effect on positive human capital. We find large effects on grade retention and high school graduation, even within pairs of identical twins. The estimates also show large effect of conduct disorder on negative human capital. Conduct disorder behaviours have a strong effect on violent and criminal behaviour since the age of 18. We also found that the effect of conduct disorder on human capital is more deleterious if these behaviours occur earlier in life.

Various robustness checks suggest that these estimates may be lower bounds of the true effect of conduct disorder on human capital. For instance, due to the routing of our survey, twins with a conduct disorder score of zero did not answer most questions on negative human capital. The size of the estimates increases if we impute that these twins did not accumulate negative human capital. In addition, the size of the estimates increases after instrumenting for measurement error in conduct disorder. Moreover, we found evidence for spill-over effects of conduct disorder within pairs of twins. These spill-over effects will bias downward the size of the within twin estimates.

Previous research (Currie and Stabile, 2006, 2007; Fletcher and Wolfe, 2008) found that conduct disorder before the age of 18 can have large deleterious longer term effects on positive and negative human capital. Our findings, which are based on a much richer set of indicators of conduct disorder and on data of twins instead of data of siblings, corroborate these results. In addition, we find that the timing of conduct disorder matters. In general, earlier occurrence of conduct disorder seems more deleterious. This indicates that early intervention is important. Previous studies have shown that early intervention programmes, like the Perry Pre-School Programme (PPP), the Syracuse Programme (SP) or the Head Start Programme (HSP) in the US, can be highly effective in ‘reducing criminal activity, promoting socioeconomic skills, and integrating disadvantaged children into mainstream society’ (Carneiro and Heckman, 2003; Schweinhart et al., 1993; Donohue and Siegelman, 1998; Lally et al., 1988; Garces et al., 2002. As such, our findings provide further evidence for the consequences of early behaviour problems and the possible gains of effective treatment early in life.

6.A Appendix

6.A.1 Variable definitions

Variable	Question
misbehaved	**L3** Did you *frequently* get into a lot of trouble with the teacher or principal for misbehaving in school? (primary or secondary school)
wagged school	**L4** Before age 18, did you ever wag school for an entire day at least twice in 1 year?
suspended/expelled	**L5** Were you ever suspended or expelled from school?
stay out late	**L6** As a child or a teenager, did you *often* stay out *much later* than you were supposed to?
sneak out at night	**L6A** Did you *often* sneak out of the house at night?
run away overnight	**L6C** Before age 18, did you ever run away from home *overnight*?
lied, used false name	**L7** Before 18, did you ever tell *a lot* of lies or use a false name or alias?
outsmarted, conned others	**L7B** Before age of 18, was there ever a period when you often outsmarted others and ŞconnedŤ them?
stole from home or family	**L8** Before age 18, did you steal money or things from your home or family more than once? If yes, did you only steal things of trivial value, like loose change or things like that?
shoplifted	**L8A** Before age 18, did you steal or shoplift from shops or other people (without their knowing) more than once? If yes, did you only steal things of trivial value like comics or lollies?
forged signature	**L8B** Before age 18, did you forge anyoneŠs signature on a cheque or credit card more than once?
damaged property	**L9** Have you ever damaged someoneŠs property *on purpose*?
started physical fights	**L10** Before age 18, did you start physical fights (with persons *other than* your brothers or sisters) 3 or more times?
used a weapon	**L11** Before age 18, did you ever use a weapon like a bat, brick, broken bottle, gun or a knife (other than in combat, when hunting, or as part of your job) to threaten or harm someone?
physically injured someone	**L12** Before age 18, (other than fighting or using a weapon) did you ever physically injure anyone on purpose?
bullied others	**L13** Before age 18, were you *often* a bully, deliberately hurting or being mean to others?
mean to animals	**L14** Before age 18, were you ever mean to animals including pets or did you hurt animals on purpose?
lighted fires	**L15** Before 18, did you ever deliberately light any fires you were not supposed to?
broke into someoneŠs car/house	**L16** Before 18, did you ever break into someoneŠs car or house or anywhere else (not because you were locked out)?
forcefully stole money or property	**L17** Before age 18, did you ever take money or property from someone else by threatening them or using force, like snatching a purse or robbing them?
forced someone into sexual activity	**L20** Before age 18, did you ever force anyone into intercourse or any other form of sexual activity?

Note: Conduct disorder statements from the TWIN89 questionnaire.

6.A.2 APA DSM-IV diagnostic criteria for conduct disorder

A. A repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate societal norms or rules are violated, as manifested by the presence of three (or more) of the following criteria in the past 12 months, with at least one criterion present in the past 6 months:
Aggression to people and animals

i. often bullies, threatens, or intimidate others
ii. often initiates physical fights
iii. has used a weapon that can cause serious physical harm to others (e.g., a bat, brick, broken bottle, knife, gun)
iv. has been physically cruel to people
v. has been physically cruel to animals
vi. has stolen while confronting a victim (e.g., mugging, purse snatching, extortion, armed robbery)
vii. has forced someone into sexual activity

Destruction of property

viii. has deliberately engaged in fire setting with the intention of causing serious damage
ix. has deliberately destroyed other's property (other than by fire setting)

Deceitfulness or theft

x. has broken into someone else's house, building, or car
xi. often lies to obtain goods or favours to avoid obligations (i.e., "cons" others)
xii. has stolen items of non-trivial value without confronting a victim (e.g., shoplifting, but without breaking and entering; forgery)

Serious violations of rules

xiii. often stays out at night despite parental prohibitions, beginning before age 13 years
xiv. has run away from home overnight at least twice while living in parental or parental surrogate home (or once without returning for a lengthy period)
xv. is often truant from school, beginning before age 13 years.

B. The disturbance in behaviour causes clinically significant impairment in social, academic, or occupational functioning.

C. If the individual is age 18 years or older, criteria are not met for antisocial personality disorder.

Specify severity:

Mild: few if any conduct problems in excess of those required to make the diagnosis and conduct problems cause only minor harms to others.

Moderate: number of conduct problems and effect on others intermediate between "mild" and "severe."

Severe: many conduct problems in excess of those required to make the diagnosis or conduct problems cause considerable harm to others.

Source: American Psychiatric Association (1994). Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington, DC: American Psychiatric Association, pp. 98-99.

Chapter 7

Conclusions

In chapter 2 we adopted different time series methodologies to investigate the impact of parole abolition and sentence reform in Virginia on reported crime rates. The Commonwealth of Virginia abolished parole and reformed sentencing for all felony offences committed on or after January 1, 1995. To examine the impact of Virginia's new legislation on reported crime rates, we considered different empirical approaches to the intervention analysis, such as regression, ARIMA, and structural time series approaches. The estimation results suggest that crimes in the eligible categories responded immediately to the passage of the new legislation in Virginia. We see not only an immediate fall in crime with the adoption of the new legislation, but also a continued decline in crime rates over several years after the passage of the new legislation. We could say that the deterrence (immediate) effect of the new legislation was about -0.1% and over the longer run as incapacitation and further deterrence kicked in, the effect of the new legislation was around -7%. When it comes to the effect of the new legislation on the particular crime offences, we find that the new legislation significantly affected burglary (-4 to -5) and murder series (-0.10). In case of rape, we find a negative significant result (-0.24) only in one of our bivariate model specifications. This result is not superior to the univariate estimation result (-0.15) which is also not found to be significantly different from zero. The possible explanation for the non-significant effect on the aggravated assault would be the very method in which this category is reported. In contrast to other crimes, which are relatively well defined, "aggravated" assault requires discretion on the part of the police taking the report to distinguish it from "simple" assault. This way of reporting leaves room for the exercise of discretion and there is a possibility that the nature of this distinction has been changing over time (Blumstein (2000)). Hence, this might be a reason for the non-significant intervention coefficient of the aggravated assault which is, by definition, violent offence.

Chapter 3 analyse macroeconomic time series of real GDP and unemployment simultaneously with burglary and theft rates for England and Wales between 1955 and 2001. By adopting univariate and multivariate structural time series approaches, we find strong evidence for cyclical behaviour in burglary and theft rates. Further, it is found that the cyclical dynamics in crime rates can be fully accounted for by cycles of economic activity. We therefore conclude that we have found no evidence of a separate crime cycle, it can be fully explained by economic cyclical behaviour. The presented analysis has accounted for various other variables that can explain variations in burglary. The explanatory variables affect more the short-term dynamics than the long-term dynamics. Further robustness checks in the analysis are carried out and they do not alter the main conclusion of this chapter. Another interesting finding of this chapter is that the model has detected the motivational and opportunity effects of crime in relation to economic factors. Although crime theory has discussed these effects, in empirical studies it has been hard to identify such effects. The methodology used in this paper has been instrumental in the ability to empirically separate these effects of GDP and unemployment on burglary and theft crimes.

Chapter 4 presents some new evidence on the effect of education on crime, looking at two different data sources from Britain. In the first, property crime convictions in England and Wales (taken from using unique Home Office Offenders Index data matched to household and employer surveys) are seen to be significantly lower amongst age cohorts where education is higher. To ensure the direction of causation runs from education to crime (and not in the opposite direction), we follow the idea of Lochner and Moretti (2004) and use exogenous changes in school leaving age laws as instruments for our education variable. Two stage least squares estimates based on this approach show that education significantly reduces property crime convictions among both men and women. Our second approach, based on using 2001 Census data on imprisonment and educational qualifications, corroborate these finding by demonstrating that having no educational qualifications significantly increases the risk of imprisonment. There is little doubt that findings from this chapter have important implications for longer-term efforts aimed at reducing crime. For example, a policy that subsidizes schooling and human capital investment is likely to reduce crime in the long run by increasing skill levels.

Chapter 5 further disentangles the strong association between human capital and crime using the Australian survey data on twins. First, we address the effects of early criminal behaviour on educational attainment. The estimates suggest that early criminal behaviour is detrimental to investment in human capital. Within pairs of twins we find that early arrests (before the age of 18) reduce educational attainment with 0.7 to 0.9 years and lower the probability of completing senior high school with 20 to 23 percentage points. In addition, the timing of the early arrest matters, such that arrests at age 13, 14 or 15 are most detrimental for educational attainment. These estimates are found after controlling for conduct disorder and early school performance. Second, we focus on the effect of human capital on crime. As early criminal activity might be an important confounder, we control for early arrests. The estimates suggest that human capital has a negative effect on crime. Completing senior high school reduces the probability of incarceration with 2 to 3 percentage points. We find similar but statistically insignificant effects on the probability of being arrested since the age of 18 and on the number of arrests. When controlling for early arrests and early behaviour problems, the estimated effect of human capital on crime reduces to less than a quarter of the previously estimated association. The strong detrimental effects of early criminal behaviour become also transparent if we consider the estimated effects of early arrests on all three measures of crime. We then find large effects of early criminal behaviour on participation in crime later on. These effects are much larger than the (isolated) impact of human capital on crime. We conclude that the strong association between human capital and crime is mainly driven by the effect of early criminal behaviour on educational attainment. Programs that succeed in preventing early criminal behaviour might yield high social and private returns.

In chapter 6 we examine the relationship between conduct disorder problems before the age of 18 and human capital accumulation, using again the Australian survey data on twins. We find large deleterious effects of conduct disorder on positive and negative human capital, even within pairs of identical twins. Conduct disorder increases the probability of grade retention and not completing high school. In addition, conduct disorder behaviours measured before the age of 18 have a strong effect on violent and criminal behaviour since the age of 18. For instance, within pairs of identical twins we find that conduct disorder reduces the probability of high school graduation with 4 to 13 percentage points and increases the probability of being arrested with 7 to 16 percentage points. Another important finding is that the effect of conduct disorder on human capital is more deleterious if these behaviours occur earlier in life. Various robustness checks, for instance sensitivity analyses that address the issues of measurement error and spill-over effects within pairs of twins, suggest that our main estimates may be lower bounds of the true effect of conduct disorder on human capital. We conclude that childhood mental health problems have high human and financial costs for families and society at large. Effective treatment early in life may yield high returns.

Summarizing, on the macro level, we find that harsher sentencing polices do reduce aggregate crime rates, but the found effects are quite small in compare to other crime reducing polices. Second, aggregate crime numbers are very responsive to the business cycle conditions in the economy, suggesting a clear economic explanation for the observed trends in crime. Third, on a micro level, we find that additional years of education among young people significantly reduce crime. If education reduces crime, then schooling will have social benefits that are not taken into account by individuals. In this case, the social returns to education may exceed the private return. However, the strong association between human capital and crime might be driven by the effect of early criminal behaviour on educational attainment. Fourth, we find large deleterious effects of conduct disorder on positive and negative human capital, where we see crime as "building capital in anti-social or criminal activities," (Currie and Stabile, 2007). We conclude that childhood mental health problems have high human and financial costs for families and society at large. Therefore, effective treatment early in life may yield high returns.

References

Agnew, R. (1991). The interactive effects of peer variables on delinquency. *Criminology 29*, 47–87.

Aigner, D. (1973). Regression with a binary independent variable subject to errors of observation. *Journal of Econometrics 1*, 49–60.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)* (4th ed.). Washington, DC: American Psychiatric Association.

Arvanites, T. M. and R. H. Defina (2006). Business cycles and street crime. *Criminology 44*(1), 139–164.

Ashenfelter, O. and A. Krueger (1994). Estimates of the economic return to schooling from a new sample of twins. *The American Economic Review 84*(5), 1157–1173.

Atkinson, A. C. (1985). *Plots, Transformations, and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. Oxford Statistical Science Series.

Atkinson, A. C., S. J. Koopman, and N. Shephard (1997). Detecting shocks: Outliers and breaks in time series. *Journal of Econometrics 80*, 387–422.

Australian Centre for Community Child Health (2006). Royal Children's Hospital Melbourne, Research Centre of the Murdoch Children Research Institute, Media Information.

Australian Institute of Criminology (2007). *Australian Crime, Facts and Figures 2007*. Canberra, Australia.

Balkin, S. and J. Keith Ord (2001). Assessing the impact of speed-limit increases on fatal interstate crashes. *Journal of Transportation and Statistics 4*(1), 1–26.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*, 169–217.

Becsi, Z. (1999). Economics and crime in the States. *Federal Reserve Bank of Atlanta, Economic Review* (First quarter).

Behrman, J. R. and M. R. Rosenzweig (2002). Does increasing women's schooling raise the schooling of the next generation. *The American Economic Review 92*(1), 323–334.

Behrman, J. R., P. Taubman, and T. J. Wales (1977). Controlling for and measuring the effects of genetics and family environment in equations for schooling and labour market success. In P. Taubman (Ed.), *Kinometrics: The Determinants of Socioeconomic Success within and between Families*, pp. 35–96. New York, NY: North-Holland Publishing Company.

Beki, C., K. Zeelenberg, and K. van Montfort (1999). An analysis of crime rates in the Netherlands 1950-1993. *British Journal of Criminology 39*, 401–415.

Benmelech, E. and C. Berrebi (2007). Attack assignments in terror organizations and the productivity of suicide bombers. *NBER Working Paper* (WP 12910).

Beveridge, S. and C. R. Nelson (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics 7*, 151–174.

Blanden, J., P. Gregg, and L. MacMillan (2006). Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *Centre for the Economics of Education Discussion Paper* (DP 73).

Blumstein, A. (2000). Disaggregating the violence trends. In A. Blumstein and J. Wallman (Eds.), *The Crime Drop in America*, pp. 13–44. New York, NY: Cambridge University Press.

Blumstein, A., J. Cohen, J. A. Roth, and C. A. Vischer (1986). *Criminal Careers and "Career Criminals"*. Washington, DC: National Academy Press.

Blumstein, A. and J. Wallman (Eds.) (2000a). *The Crime Drop in America.* New York, NY: Cambridge University Press.

Blumstein, A. and J. Wallman (2000b). The recent rise and fall of American violence. In A. Blumstein and J. Wallman (Eds.), *The Crime Drop in America*, pp. 1–12. New York, NY: Cambridge University Press.

Boggess, S. and J. Bound (1993). Did criminal activity increase during the 1980s? Comparison across data sources. *NBER Working Paper Series* (WP 4431).

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008). The economics and psychology of personality traits. *Journal of Human Resources 43*(4).

Boswijk, H. P., M. H. C. Kok, and M. J. van Leeuwen (2002). Naar een gecombineerd VEC-model voor jeugd- en volwassenencriminaliteit. *SEO and WODC Report.*

Bound, J. and G. Solon (1999). Double trouble: on the value of twins-based estimation of the return to schooling. *Economics of Education Review 18*, 169–182.

Box, G. E. P. and G. M. Jenkins (1976). *Time Series Analysis: Forecasting and Control.* San Francisco, CA: Holden-Day.

Box, G. E. P. and G. C. Tiao (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association 70*(349), 70–79.

Box, S. (1987). *Recession, crime and punsihment.* London: Palgrave Macmillan.

Britt, C. L. (1997). Reconsidering the unemployment and crime relationship: Variations by age group and historical period. *Journal of Quantitative Criminology 13*(4), 405–407.

Britt, C. L. (2001). Testing theory and the analysis of time series data. *Journal of Quantitative Criminology 17*(4), 343–357.

Burdett, K., R. Lagos, and R. Wright (2003). Crime, inequality, and unemployment. *The American Economic Review 93*(5), 1764–1777.

Calvó-Armengol, A. and Y. Zenou (2004). Social networks and crime decision: The role of social structure in facilitating delinquent behaviour. *International Economic Review 45*(3), 939–958.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press.

Cantor, D. and K. C. Land (1985). Unemployment and crime rates in the post World War II United States: A theoretical and empirical analysis. *American Sociological Review 50*(3), 317–332.

Cantor, D. and K. C. Land (2001). Unemployment and crime rate fluctuations: A comment on Greenberg. *Journal of Quantitative Criminology 17*(4), 329–342.

Cappell, C. L. and G. Sykes (1991). Prison commitments, crime and unemployment: A theoretical and empirical specification for the United States. *Journal of Quantitative Criminology 17*(2), 155–199.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *The Handbook of Labour Economics*, Volume 3A, Chapter 30. North-Holland: Elsevier.

Card, D. (2001). Estimating the returns to schooling: Progress on some persistent econometric problems. *Econometrics 69*(5), 1127–1160.

Carneiro, P. and J. J. Heckman (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, and B. M. Friedman (Eds.), *Inequality in America: What Role for Human Capital Policies?*, pp. 77–239. Cambridge, MA: MIT Press.

Case, A. C. and L. F. Katz (1991). The company you keep: The effects of family and neighbourhood disadvantaged youths. *NBER Working Paper* (WP 3705).

Chan, S.-M., P. East, S. Ali, and M. Neophytou (2002). Primary and secondary education in England and Wales: From 1944 to the present day. Technical Report.

Charlton, K. and M. McCall (2004). Statistics on juvenile detention in Australia: 1981-2003. Australian Institute of Criminology.

Chevalier, A. (2004). Parental education and child's education: A natural experiment. Working Paper.

Chevalier, A., C. Harmon, V. O'Sullivan, and I. Walker (2005). The impact of parental income and education on the schooling of their children. *The Institute for Fiscal Studies* (WP05/05).

Chiricos, T. (1987). Rates of crime and unemployment: An analysis of aggregate research evidence. *Social Problems 34*(2), 187–212.

Clark, J., J. Austin, and D. A. Henry (1997). Three strikes and you're out: A review of state legislation. *Office of Justice Statistics NCJ 165369*. U.S. Department of Justice.

Clark, P. (1987). The cyclical component of U.S. economic activity. *The Quarterly Journal of Economics 102*, 797–814.

Clarke, R. V. and P. M. Harris (1992). Auto theft and its prevention. *Crime and Justice 16*, 1–54.

Commandeur, J. J. F. and S. J. Koopman (2007). *An Introduction to State Space Time Series Analysis*. Oxford: Oxford University Press.

Cook, P. J. and G. A. Zarkin (1985). Crime and the business cycle. *Journal of Legal Studies 14*, 115–128.

Corman, H., T. Joyce, and N. Lovitch (1987). Crime, deterrence, and the business cycle in New York: A VAR approach. *The Review of Economics and Statistics 69*(4), 695–700.

Currie, J. and M. Stabile (2006). Child mental health and human capital accumulation: The case of ADHD. *Journal of Health Economics 25*(6), 1094–1118.

Currie, J. and M. Stabile (2007). Mental health in childhood and human capital. *NBER Working Paper* (WP 13217).

Deadman, D. F. (2003). Forecasting residential burglary. *International Journal of Forecasting 19*, 567–578.

Dhiri, S., S. Brand, R. Harries, and R. Price (1999). Modelling and predicting property crime trends. *Home Office Research Study* (198).

DiIulio, Jr., J. J. (1996). Help wanted: Economists, crime and public policy. *Journal of Economic Perspectives 10*(1), 3–24.

Donohue, J. and P. Siegelman (1998). Allocating resources among prisons and social programs in the battle against crime. *Journal of Legal Studies 27*, 1–43.

Doyle, O., C. P. Harmon, and I. Walker (2007). The impact of parental income and education on child health: Further evidence for England. *UCD Geary Institute Discussion Paper Series* (Geary WP/6/2007).

Draca, M., S. Machin, and R. Witt (2008). Panic on the streets of London: Police, crime and the July 2005 terror attacks. *CEP Discussion Paper* (No. 852).

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

Eck, J. and E. Maguire (2000). Have changes in policing reduced violent crime? An assessment of the evidence. In A. Blumstein and J. Wallman (Eds.), *The Crime Drop in America*, pp. 207–265. New York, NY: Cambridge University Press.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy 81*(3), 521–565.

Ehrlich, I. (1975). *On the Relation between Education and Crime*, Chapter 12. New York, NY: McGraw-Hill Book Co.

Eide, E. (2000). Economics of criminal behaviour. In B. Bouckaert and G. D. Geest (Eds.), *Encyclopaedia of Law and Economics*, Volume 5, pp. 345–389. Cheltenham: Edward Elgar.

Emery, R. E. and L. Lauman-Billings (1998). An overview of the nature, causes, and consequences of abusive family relationships: Toward differentiating maltreatment and violence. *American Psychologist 53*, 121–135.

Entorf, H. and H. Spengler (2000). Socioeconomic and demographic factors of crime in germany: Evidence from panel data of the German states. *International Review of Law and Economics 20*, 75–106.

Farmer, E. M. Z. (1995). Extremity of externalizing behaviour and young adult outcomes. *Journal of Child Psychology 36*(4), 617–632.

Farrington, D. (1973). Self-reports of deviant behaviour: Predictive and stable? *Journal of Criminal Law and Criminology 64*(1), 99–110.

Farrington, D. P. (1986). *Age and Crime*. Chicago: University of Chicago Press.

Farrington, D. P. (2001). *Predicting Persistent Young Offenders*. United Kingdom: Macmillan Press Ltd.

Farrington, D. P., G. C. Barnes, and S. Lambert (1996). The concentration of offending in families. *Legal and Criminological Psychology 1*, 47–63.

Farrington, D. P., S. Lambert, and D. J. West (1998). Criminal careers in two generations of family members in the Cambridge study of delinquent development. *Studies on Crime and Crime Prevention 7*, 1–22.

Feinstein, L. (2002). Quantitative estimates of the social benefits of learning, 1: Crime. *Wider Benefits of Learning Research Report* (5).

Fisher, D. G., G. L. Reynolds, M. M. Wood, and M. E. Johnson (2004). Reliability of arrest and incarceration questions on the risk behaviour assessment. *Crime and Delinquency 50*(1), 24–31.

Fisher, F. and D. Nagin (1978). On the feasibility of identifying the crime function in a simultaneous model of crime rates and sanction levels. In A. B. et. al. (Ed.), *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, pp. 361–399. Washington, DC: National Academy of Sciences.

Fletcher, J. and B. L. Wolfe (2008). Child mental health and human capital accumulation: The case of ADHD revisited. *Journal of Health Economics 27*, 794–800.

Fougère, D., F. Kramarz, and J. Pouget (2007). Youth unemployment and crime in france. *CREST-INSEE Working Paper* (33).

Fox, J. A. (2000). Demographics and U.S. homicide. In A. Blumstein and J. Wallman (Eds.), *The Crime Drop in America*, pp. 288–317. New York, NY: Cambridge University Press.

Freeman, R. (1983). Crime and the labour market. In J. Q. Wilson (Ed.), *Crime and Public Policy*, pp. 89–106. San Francisco: ICS Press.

Freeman, R. B. (1996). Why do so many young American men commit crimes and what might we do about it? *Journal of Economic Perspectives 10*(1), 25–42.

Funk, P. and P. Kugler (2003a). Dynamic interactions between crimes. *Economic Letters 79*, 291–298.

Funk, P. and P. Kugler (2003b). Identifying efficient crime-combating policies by VAR models: The example of Switzerland. *Contemporary Economic Policy 21*(4), 525–538.

Galindo-Rueda, F. (2003). The intergenerational effect of parental schooling: Evidence from the British 1947 school leaving age reform. Working Paper.

Garces, E., D. Thomas, and J. Currie (2002). Longer-term effects of Head Start. *American Economic Review 92*(4), 999–1012.

Gershoff, E. T. (2002). Parental corporal punishment and associated child behaviours and experiences: A meta-analytic and theoretic review. *Psychological Bulletin 128*, 539–579.

Gottfredson, M. and T. Hirschi (1986). The true value of lambda would appear to be zero: An essay on career criminals, criminal careers, selective incapacitation, cohort studies, and related topics. *Criminology 24*, 213–234.

Gould, D., D. Mustard, and B. Weinberg (2002). Crime rates and local labour market opportunities in the United States: 1979-1997. *Review of Economics and Statistics 84*, 45–61.

Granger, C. W. J. and P. Newbold (1976). Forecasting transformed series. *Journal of Royal Statistical Society B 38*, 189–203.

Greenberg, D. F. (2001). Time series analysis of crime rates. *Journal of Quantitative Criminology 17*(4), 291–327.

Gregg, P. and S. Machin (2000). Child development and success or failure in the youth labour market. In D. G. Blanchflower and R. B. Freeman (Eds.), *NBER Comparative Labour Market Series*, pp. 247–288. Chicago: The University of Chicago Press.

Griliches, Z. (1979). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy 87*(5), 37–64.

Grogger, J. (1991). Certainty vs. severity of punishment. *Economic Inquiry 29*(2), 297Ű309.

Grogger, J. (1998). Market wages and youth crime. *Journal of Labour Economics 16*(4), 756–791.

Grogger, J. (2000). An economic model of recent trends in violence. In A. Blumstein and J. Wallman (Eds.), *The Crime Drop in America*, pp. 266–287. New York, NY: Cambridge University Press.

Hale, C. and D. Sabbagh (1991). Testing the relationship between unemployment and crime: A methodological comment and empirical analysis using time series data from England and Wales. *Journal of Resident Delinquency 28*, 400–417.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.

Hansen, K. (2003). Education and the crime-age profile. *British Journal of Criminology 43*, 141–168.

Hardt, R. H. and S. Hardt-Petersen (1977). On determining the quality of the delinquency self-report method. *Journal of Research in Crime and Delinquency 14*, 247–261.

Harmon, C. and I. Walker (1995). Estimates of the economic return to schooling for the United Kingdom. *The American Economic Review 85*(5), 1278–1286.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Harvey, A. C. (1993). *Time Series Models*. Great Britain: Financial Times Prentice Hall.

Harvey, A. C. (1996). Intervention analysis with control groups. *International Statistical Review 64*(3), 313–328.

Harvey, A. C. and J. Durbin (1986). The effects of seat belt legislation on British road casualties: A case study in structural time series modelling. *Journal of the Royal Statistical Society, Series A 149*, 187–227.

Harvey, A. C. and C. Fernandez (1989). Time series models for count or qualitative observations. *Journal of Business & Economic Statistics 7*(4), 407–417.

Harvey, A. C. and S. J. Koopman (1992). Diagnostic checking of unobserved components time series models. *Journal of Business and Economic Statistics 10*(4), 377–389.

Harvey, A. C. and S. J. Koopman (1997). Multivariate structural time series models. In C. Heij, H. Schumacher, B. Hanzon, and C. Praagman (Eds.), *Systematic dynamics in economic and financial models*, pp. 269–98. Chichester: John Wiley and Sons.

Heckman, J. J. and D. V. Masterov (2007). The productivity argument for investing in young children. *IZA Discussion Paper* (2725).

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labour market outcomes and social behaviour. *Journal of Labour Economics 24*(3), 411–481.

Hendry, D. F. and J. A. Doornik (1996). *Empirical Econometric Modelling Using Pc-Give 9.0 for Windows*. London.

Hindelang, M. J., T. Hirschi, and J. G. Weiss (1981). *Measuring Delinquency*. Beverly Hills, CA: Sage.

Hird, C. and C. Ruparel (2007). Seasonality in recorded crime: Preliminary findings. *Home Office Online Report* (02/07).

Hirschi, T. and M. Gottfredson (1983). Age and the explanation of crime. *American Journal of Sociology 89*, 552–584.

Hirschi, T. and M. Gottfredson (1995). Control theory and the life-course perspective. *Studies on Crime and Crime Prevention 4*, 131–142.

Home Office (1995). Criminal careers of those born between 1953-1973. *Home Office Statistical Bulletin*. London: Home Office.

Horney, J. and I. Marshall (1992). An experimental comparison of two self-report methods for measuring lambda. *Journal of Research in Crime and Delinquency 29*(1).

Huizinga, D. and D. S. Elliott (1986). Reassessing the reliability and validity of self-report delinquency measures. *Journal of Quantitative Criminology 2*(4), 293–327.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Johnson, B., A. Golub, and E. Dunlap (2000). The rise and decline of hard drugs, drug markets, and violence in inner-city New York. In A. Blumstein and J. Wallman (Eds.), *The Crime Drop in America*, pp. 164–206. New York, NY: Cambridge University Press.

Johnston, J. and J. DiNardo (1997). *Econometric Methods*. New York, NY: The McGraw-Hill Companies, Inc.

Jovičić, M. (1989). *Ekonometrijski metodi*. Beograd: Ekonomski Fakultet Univerziteta u Beogradu.

Kane, T. J., C. E. Rouse, and D. Staiger (1999). Estimating returns to schooling when schooling is misreported. *Working paper* (419). Princeton University.

Kessler, D. and S. D. Levitt (1999). Using sentence enhancements to distinguish between deterrence and incapacitation. *Journal of Law and Economics 42*, 343–363.

Koopman, S. J., A. C. Harvey, J. A. Doornik, and N. Shephard (2006). *STAMP 7.0: Structural Time Series Analyser, Modeller and Predictor*. London.

Koopman, S. J., A. Lucas, M. Ooms, K. van Montfort, and V. van der Geest (2008). Estimating systematic continuous-time trends in recidivism using a non-Gaussian panel data model. *Statistica Neerlandica 62*(1), 104–130.

Krueger, A. B. and J. Malečková (2003). Education, poverty and terrorism: Is there a causal connection? *Journal of Economic Perspectives 17*(4), 119–144.

Lally, J. R., P. L. Mangione, and A. S. Honig (1988). The Syracuse university family development research programme: Long-range impact of an early intervention with low-income children and their families. In D. Powell (Ed.), *Parent Education as Early Childhood Intervention: Emerging directions in theory, research, and practice*, pp. 79–104. Norwood, NJ: Ablex.

Landau, S. F. and D. Fridman (1993). The seasonality of violent crime: The case of robbery and homicide in Israel. *Journal of Research in Crime and Delinquency 30*(2), 163–191.

Layard, R. (2005). *Happiness: Lessons from a New Science*. London: Penguin Books Ltd.

Le, A. T., P. W. Miller, A. C. Heath, and N. Martin (2005). Early childhood behaviours, schooling, and labour market outcomes: Estimates from a sample of twins. *Economics of Education Review 24*, 1–17.

Levitt, S. D. (1997). Using electoral cycles in police hiring to estimate the effect of police on crime. *The American Economic Review 87*(3), 270–290.

Levitt, S. D. (2001). Alternative strategies for identifying the link between unemployment and crime. *Journal of Quantitative Criminology 17*(4), 377–390.

Levitt, S. D. and S. J. Dubner (2005). *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. London: Penguin Books Ltd.

Levitt, S. D. and L. Lochner (2001). *The determinants of juvenile crime*, pp. 327–373. Chicago, IL: University of Chicago Press.

Lindeboom, M., A. Llena-Nozal, and B. van der Klauw (2008). Parental education and child health: Evidence from a schooling reform. *Journal of Health Economics* (Forthcoming).

Lochner, L. (2004). Education, work, and crime: A human capital approach. *International Economic Review 45*, 811–843.

Lochner, L. (2007). Education and crime. Working Paper.

Lochner, L. and E. Moretti (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *The American Economic Review 94*, 155–189.

Loeber, R. (1982). The stability of antisocial and delinquent child behaviour: A review. *Child Development 53*, 1431–1446.

Loftin, C., M. Heumann, and D. McDowall (1983). Mandatory sentencing and firearms violence: Evaluating an alternative to gun control. *Law & Society Review 17*(2), 287–318.

Ludwig, J., G. J. Duncan, and P. Hirschfield (2001). Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment. *Quarterly Journal of Economics 116*, 655–679.

Lundborg, P. (2008). The health returns to education - What can we learn from twins? Working Paper.

Lynch, S. K., E. Turkheimer, B. M. D'Onofrio, J. Mendle, R. E. Emery, W. S. Slutske, and N. G. Martin (2006). A genetically informed study of the association between harsh punishment and offspring behavioural problems. *Journal of Family Psychology 20*(2), 190–198.

Machin, S. and C. Meghir (2004). Crime and economics incentives. *Journal of Human Resources 39*(4), 958–79.

Machin, S. and S. Vujić (2008). Crime and education: New evidence from Britain. CEP Discussion Paper.

Maddux, J. and D. Desmond (1975). Reliability and validity of information from chronic heroin users. *Journal of Psychiatric Research 12*(2), 87–95.

Marvell, T. B. and C. E. Moody (1994). Prison population growth and crime reduction. *Journal of Quantitative Criminology 10*(2), 109–140.

Marvell, T. B. and C. E. Moody (1996). Specification problems, police levels, and crime rates. *Criminology 34*(4), 609–646.

McCleary, R. and R. A. Hay, Jr. (1980). *Applied Time Series Analysis for the Social Sciences*. Beverly Hills and London: SAGE Publications, Inc.

McDowall, D., C. Loftin, and B. Wiersema (1995). Easing concealed firearm laws: Effects on homicide in three states. *Journal of Criminal Law and Criminology (Northwestern) 86*(1), 193–206.

McDowall, D., R. McCleary, E. E. Meidinger, and R. A. Hay, Jr. (1980). *Interrupted Time Series Analysis*. SAGE Publications, Inc.

Mieczkowski, T. (1990). *The accuracy of self-reported drug use: An evaluation and analysis of new data*, pp. 275–302. Cincinnati: Anderson Publishing.

Miller, P., C. Mulvey, , and N. Martin (2006). The return to schooling: Estimates from a sample of young Australian twins. *Labour Economics 13*, 571–587.

Miller, P., C. Mulvey, and N. Martin (1995). What do twins studies reveal about the economic returns to education? A comparison of Australian and U.S. findings. *The American Economic Review 85*(3), 586–599.

Moolenaar, D., F. van Tulder, and G. Huijbregts (2002). Trends in Dutch juvenile delinquency, law enforcement and justice. *WODC Report*.

Nieves, K., J. Draine, and P. Solomon (2000). The validity of self-reported criminal arrest history among clients of a psychiatric probation and parole service. *Journal of Offender Rehabilitation 30*(3/4).

Nyblom, J. and A. C. Harvey (2001). Testing against smooth stochastic trends. *Journal of Applied Econometrics 16*(3), 415–429.

O'Brien, R. M. (1999). Measuring the convergence/divergence of "serious crime" arrest rates for males and females: 1960-1995. *Journal of Quantitative Criminology 15*, 97–114.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *The American Economic Review 96*(1), 152–175.

Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics 91*(11-12), 2213–2229.

Orwin, R. G. (1997). Twenty-one years old and counting: The interrupted time series comes of age. In E. Chelimsky and W. R. Shadish (Eds.), *Evaluation for the 21st Century: A Handbook*, pp. 443–465. Thousand Oaks, CA: Sage Publications, Inc.

Papps, K. and R. Winkelmann (1998). Unemployment and crime: New answers to an old question. *IZA Discussion Paper* (25).

Parke, R. D. (1979). *Punishment: Effects, side effects, and alternative strategies*, pp. 71–99. New York, NY: Academic Press.

Paternoster, R. and S. D. Bushway (2001). Theoretical and empirical work on the relationship between unemployment and crime. *Journal of Quantitative Criminology 17*(4), 391–407.

Pyle, D. J. and D. F. Deadman (1994). Crime and the business cycle in post-war Britain. *British Journal of Criminology 34*(3), 339–357.

Ralston, R. W. (1999). Economy and race: Interactive determinants of property crime in the United States, 1958-1995: Reflections on the supply of property crime. *American Journal of Economics and Sociology 58*(3), 405–434.

Raphael, S. and R. Winter-Ember (2001). Identifying the effect of unemployment on crime. *Journal of Law and Economics 44*(1), 259–283.

Reilly, B. and R. Witt (1996). Crime, deterrence and unemployment in England and Wales: An empirical analysis. *Bulletin of Economic Research 48*(2), 137–159.

Reiss, A. J. (1988). *Co-offending and criminal careers*, Volume 10, pp. 117–170. Chicago: University of Chicago Press.

Roberts, D. (2005). Virginia incident-based crime reporting. *Virginia Criminal Justice Research Centre Research Bulletin*.

Russell, C. and B. Miller (1983). *Profile of a terrorist*, pp. 45–60. Wilmington, Del.: Scholarly Resources Inc.

Rutter, M., H. Giller, and A. Hagell (1998). *Antisocial Behaviour by Young People*. Cambridge: Cambridge University Press.

Saridakis, G. (2004). Violent crime in the United States of America: A time-series analysis between 1960-2000. *European Journal of Law and Economics 18*, 203–221.

Schweinhart, L. J., H. V. Barnes, and D. P. Weikart (1993). *Significant Benefits: The High-Scope Perry Pre-School Study through Age 27*. Ypsilanti, MI: High Scope Press.

Searight, H. R., F. Rottnek, and S. L. Abby (2001). Conduct disorder: Diagnosis and treatment in primary care. *American Family Physician 63*(8), 1579–1588.

Short, Jr., J. F. and F. I. Nye (1957). Reported behaviour as a criterion of deviant behaviour. *Social Problems 5*(3), 207–213.

Slade, E. P. and L. S. Wissow (2007). The influence of childhood maltreatment on adolescents' academic performance. *Economics of Education Review 26*(5), 604–614.

Small, J. and C. Lewis (1996). Economic crime in New Zealand: Causation or coincidence? *University of Auckland Working Paper* (158).

Smith, C. A. and D. P. Farrington (2004). Continuities in antisocial behaviour and parenting across three generations. *Journal of Child Psychology and Psychiatry 45*, 230–247.

Smith, M. D., J. Devine, and J. Sheley (1992). Crime and unemployment: Effects across age and race categories. *Sociological Perspectives 35*(4), 551–572.

Soothill, K., B. Francis, B. Sanderson, and E. Ackerley (2000). Sex offenders: Specialists, generalists – or both. *British Journal of Criminology 40*(1), 56–67.

Spear, L. P. (2000). The adolescent brain and age-related behavioural manifestations. *Neuroscience and Biobehavioural Reviews 24*, 417–463.

Spelman, W. (2000). The limited importance of prison expansion. In A. Blumstein and J. Wallman (Eds.), *The Crime Drop in America*, pp. 97–129. New York, NY: Cambridge University Press.

Staiger, D. and J. H. Stock (1997). Instrumental variables with weak instruments. *Econometrica 65*(3), 557–586.

Stata (2006). *Statistical Software: Release 9.0*. College Station, TX: Stata Corporation.

Stock, J. and M. W. Watson (1988). Testing for common trends. *Journal of American Statistical Association 3*(4), 1097–1107.

Straus, M. A. (1994). *Beating the devil out of them: Corporal punishment in American families*. New York, NY: Lexington Books.

Swaine Thomas, D. (1927). *Social Aspects of Business Cycles*. New York, NY: Knopf.

Tauchen, H., A. D. Witte, and H. Griesinger (1994). Criminal deterrence: Revisiting the issue with a birth cohort. *Review of Economics and Statistics 76*, 399–412.

Thornberry, T. P. and R. L. Christenson (1984). Unemployment and criminal involvement: An investigation of reciprocal causal structures. *American Sociological Review 49*(3), 398–411.

Thornberry, T. P. and M. D. Krohn (1997). Peers, drug use, and delinquency. In D. Stoff, J. Breiling, and J. D. Maser (Eds.), *Handbook of antisocial behaviour*, pp. 218–233. New York, NY: Wiley.

Thornberry, T. P. and M. D. Krohn (2000). The self-report method for measuring delinquency and crime. *Criminal Justice 4*, 33–83.

Verbeek, M. (2000). *A Guide to Modern Econometrics*. Chichester: John Wiley & Sons, Ltd.

Vollaard, B. and P. Koning (2006). Estimating police effectiveness with individual victimisation data. *CPB Discussion Paper* (47).

Vujić, S. (2003). VAR approach to modelling causes of property crime. Final paper for the Econometrics Summer School in Copenhagen.

Vujić, S. and S. J. Koopman (2008). Intervention time series analysis of crime rates: The impact of sentence reforms in Virginia. Working Paper.

Vujić, S., D. Webbink, P. Koning, and N. Martin (2008). The effect of childhood conduct disorder on human capital: Evidence from a cohort of young Australian twins. CPB Discussion Paper.

Weiss, J. G. (1986). Issues in the measurement of criminal careers. In A. Blumstein, J. Cohen, J. A. Roth, and C. A. Visher (Eds.), *Criminal Careers and "Career Criminals"*, Volume 2, pp. 1–51. Washington, D.C.: National Academy Press.

Weiss, R. D., L. M. Najavits, S. F. Greenfield, J. A. Soto, S. R. Shaw, and D. Wyner (1998). Validity of substance use self-reports in dually diagnosed outpatients. *American Journal of Psychiatry* (155), 127–128.

Wellington, J., A. M. Bathmaker, C. Hunt, G. McCulloch, and P. Sikes (2005). *Succeeding With Your Doctorate*. London: SAGE Publications Ltd.

Wilson, J. Q. and R. J. Herrnstein (1985). *Crime and Human Nature*. New York, NY: Simon and Schuster.

Witt, R. and A. Dryden Witte (1998). Crime, imprisonment, and female labour force participation: A time-series approach. *NBER Working Paper Series* (WP 6786).

Wolpin, K. I. (1978). An economic analysis of crime and punishment in England and Wales, 1894-1967. *Journal of Political Economy 815.*

Samenvatting (Summary in Dutch)

De criminaliteit is in brede zin een van de grootste sociale en politieke problemen in vele landen. Het is een onderwerp waarover veel gediscussieerd wordt onder beleidsmakers en maatschappelijke groeperingen. Het is ook een favoriet onderwerp voor politici tijdens verkiezingscampagnes. Een hoge criminaliteit resulteert in enorme maatschappelijke kosten die men veelal niet kan terugvinden in de officiële cijfers over welzijn en geluk van landen. De wetenschappelijke en invloedrijke studies van Garry Becker (1968) en Isaac Ehrlich (1973) hebben aangetoond dat het niveau van de criminaliteit afhangt van economische factoren en prikkels. Deze inzichten zijn tegenwoordig gemeengoed onder economen. De standaard in economische modellen voor crimineel gedrag bepaalt dat het individueel niveau van criminele activiteiten afhangt van het inkomen dat uit legale activiteiten kan worden verkregen, van persoonlijke voorkeuren en van andere individuele en exogene factoren. Ondanks veel en uitgebreid wetenschappelijk onderzoek op dit gebied, is het algemene publiek veelal niet doordrongen van de economische aspecten van criminaliteit en de implicaties ervan. Dit proefschrift presenteert een aantal wetenschappelijk studies die deze aspecten en implicaties verder onder het licht zal brengen.

Vanuit een economische perspectief kan het niveau van de criminaliteit door een vraag en aanbod model worden beschreven. In een vraag en aanbod model bieden criminelen het niveau van de criminaliteit, vraagt het publiek bescherming tegen de criminaliteit en biedt de overheid publieke bescherming. Het aanbod van de criminaliteit wordt gezien als een keuze tussen legale werkzaamheden aan de ene kant en criminele activiteiten aan de andere kant. De keuze zelf hangt af van de netto-baten van crimineel gedrag. Dit zijn de baten van criminele activiteiten die uitstijgen boven alle kosten die in relatie tot criminele activiteiten moeten worden gemaakt. Deze baten hangen vooral af van het type misdaad en van de crimineel zelf. Het kunnen zowel monetaire als psychische baten zijn, zoals bijvoorbeeld de spanning van het gevaar, de acceptatie onder leeftijdsgenoten, het gevoel iets volbracht te hebben of puur lichamelijke tevredenheid in het geval van verkrachting. De kosten van criminele activiteiten kunnen ingedeeld worden in directe en indirecte (opportuniteits-) kosten. Directe kosten van de criminaliteit zijn alle materiële kosten (uitrusting, geweer, voertuigen), psychische kosten (schuld, angst, risico) en verwachte strafkosten (boetes, vrijheidsstraf, andere straffen, kosten van een juridisch proces). Opportuniteitskosten van de misdaad zijn alle gemiste netto-baten (baten minus de kosten) van de legale activiteiten tijdens het plannen, uitvoeren en verbergen van de criminele activiteit. Onderwijs verhoogt bijvoorbeeld de opportuniteitskosten van misdaad omdat het legale inkomen door onderwijs zal stijgen. Het aanbod van de criminaliteit heeft een positieve relatie met de netto-baten van criminele activiteiten. Dit betekent dat criminelen hun activiteiten verhogen als de netto-baten hoger zijn. Voorbeelden van de condities die het aanbod van de criminaliteit doen toenemen zijn demografie (een grotere proportie van jeugd), minder werkgelegenheid en vermindering van vrijheidsstraffen. Terwijl vermogensdelicten het meest gevoelig zijn voor economische factoren, zijn veel geweldsdelicten het bijproduct van de criminele activiteiten voor economische (financiële) baten en kunnen derhalve eveneens door economische factoren worden uitgelegd.

Dit proefschrift richt zich voornamelijk op het aanbod aspect van de criminaliteit. De relatie tussen criminaliteit en economische factoren is geanalyseerd op basis van geaggregeerde (nationale) tijdreeksen (voor de VS en het Verenigd Koninkrijk), geaggregeerde cross-sectionele bestanden (voor het Verenigd Koninkrijk) en individuele onderzoeksgegevens verkregen via surveys (voor Australië). Het doel van dit proefschrift is om empirisch te onderzoeken of we criminaliteit als rationeel gedrag kunnen beschouwen. Als dit laatste het geval is, dan kan economisch en sociaal beleid de belangrijkste instrumenten leveren om criminaliteit te voorkomen en te bestrijden.

Hoofdstuk 2 bestudeert de trends in de criminaliteitscijfers van Virginia (VS). Historisch gezien zijn deze criminaliteitscijfers toegenomen in de periode 1960-1970 terwijl ze in alle categorieën dramatisch zijn gedaald tijdens de jaren 1990. Levitt (2005) schrijft dat de zestiger en zeventiger jaren “the heyday of a liberal justice system and the criminal’s right movement” waren. De kans om gestraft te worden was zo laag dat het lonend was om crimineel te worden. In de jaren 1980 is deze liberale houding ten opzichte van crimineel gedrag dramatisch veranderd. De rechten van verdachten en veroordeelden werden ingeperkt en er werden strenge straffen geïntroduceerd. Tijdens de negentiger jaren namen bijna alle staten van de VS een of meerdere wetten aan met verplichte straffen tot gevolg. Veel van deze wetten hadden gemeen dat de vrijheidsstraf werd verhoogd bij individuen die al eerder tot een gevangenisstraf waren veroordeeld. Er bestaat nog steeds discussie tussen criminologen en beleidsmakers over welke factoren het meest hebben bijgedragen aan de enorme daling in de jaren negentig van de criminaliteitscijfers in de VS. De belangrijkste factoren zijn wellicht de legalisatie van abortus in de jaren 1970, verbeterde economische omstandigheden, lokale beleidsveranderingen, veranderingen in de marktomstandigheden van crack en andere drugs, hogere arrestatiekansen met gevangenisstraf tot gevolg, et cetera. Dit hoofdstuk bestudeert de statistische relatie tussen inspanningen om de criminaliteit te bestrijden en de daling van de criminaliteit door de tijd, op basis van tijdreeksgegevens uit Virginia (VS) in de periode 1984-2006.

We bestuderen het effect van het nieuwe strafrecht in Virginia op het aantal gerapporteerde vermogensdelicten (diefstal met inbraak, diefstal van motorvoertuigen en diefstal met geweld) en geweldsdelicten (zware mishandeling, moord en verkrachting) met behulp van verschillende tijdreeks modellen. Virginia schafte het “parole” af en hervormde de vrijheidsstraf voor alle ernstige misdaden uitgevoerd op of na 1 januari 1995. Het netto effect van de nieuwe wet was een omvangrijke verhoging van straffen voor alle geweldsdelicten (maar vooral voor moord en verkrachting) en ook voor misdadigers die al eerder geweldsdelicten hadden gepleegd. Wij hebben het effect van de nieuwe strafmaatregelen in Virginia op de criminaliteitscijfers bestudeerd aan de hand van verschillende tijdreeks modellen. Specifieke methoden uit interventie analyse zijn gebruikt. Het effect op de criminaliteitscijfers is gemodelleerd als een *stap interventie*. De bestaande empirische literatuur geeft weinig steun aan de effecten op het afgeschafte parool en de hervormde vrijheidsstraffen in Virginia. Kessler en Levitt (1999) melden dat eerder onderzoek naar het effect van de verhoogde vrijheidsstraf aan de gerapporteerde delicten in de VS geen duidelijke resultaten geeft en vaak geen effect vindt. In dit hoofdstuk bevestigen we deze conclusies op basis van de univariate en multivariate tijdreeks modellen. In het geval van multivariate tijdreeks modellen, laten we controle groepen in de analyse toe. Wij illustreren de flexibiliteit van deze multivariate aanpak en tonen aan dat er effecten op criminaliteitscijfers gemeten kunnen worden.

Hoofdstuk 3 richt zich op misdaad trends in het Verenigd Koninkrijk. We kijken vooral naar de effecten van de "business cycle" op vermogensdelicten (diefstal en diefstal met inbraak) in Engeland en Wales tijdens de tweede helft van de twintigste eeuw. In het Verenigd Koninkrijk is de algemene criminaliteit snel toegenomen in de jaren zeventig en tachtig. Aan het begin van de jaren negentig bereikte de criminaliteit het hoogste niveau en is een algemene publieke discussie geworden. De toename van de criminaliteit was veel groter dan in de rest van Europa. Bovendien zijn de vermogensdelicten in het Verenigd Koninkrijk tijdens de jaren 1970-1990 snel gegroeid. Machin en Meghir (2004) laten in hun studie zien dat het niveau van de vermogensdelicten in de jaren negentig in Engeland en Wales hoger was dan in de VS.

De "business cycle" heeft een groot effect op de structuur van de economische dynamiek en daardoor op het gedrag van mensen. Dit effect is waargenomen bij diverse sociale indicatoren zoals het aantal schoolinschrijvingen, de geboortecijfers en de arbeidsmarktparticipatie. Het is dan ook geen verrassing dat de criminaliteitscijfers ook door de "business cycle" beïnvloed worden. De literatuur suggereert dat vermogensdelicten een contra-cyclisch patroon hebben: dalingen en stijgingen in deze delicten komen overeen met de opgaande en neergaande bewegingen van economische activiteiten. Aan de andere kant worden de geweldsdelicten minder beïnvloed door cyclische bewegingen in de economie (wellicht met uitzondering van diefstal met geweld omdat dit delict duidelijk door profijt wordt gemotiveerd).

Om economische motieven van diefstal en diefstal met geweld in de periode van 1955 tot 2001 in Engeland en Wales te bestuderen, gebruiken wij zowel univariate als multivariate structurele tijdreeks modellen. Wij modelleren de "business cycle" als een stationair *autoregressive moving average* (ARMA) proces. Dit is een nieuwe aanpak in vergelijking met het bestaande onderzoek naar het effect van de "business cycle" in de literatuur over criminaliteit. In het multivariate structurele tijdreeks model worden tijdreeksen van diefstal en diefstal met inbraak simultaan geanalyseerd met tijdreeksen van reële BNP, werkeloosheid en politiecijfers. Dit brengt ons bij de tweede bijdrage van dit hoofdstuk – de gelijktijdige analyse van de criminele (diefstal en diefstal met inbraak), economische (BNP en werkloosheid) en afschrikking (politie) tijdreeksen. Wij controleren ook voor andere verklarende variabelen die belangrijk zijn voor de verklaring van de variaties in de criminaliteit. Voorbeelden van verklarende variabelen zijn strafrecht variabelen (veroordeling, straflengte en vrijheidsstraf), demografische variabelen (geslacht, jeugd tussen de leeftijd van 15 en 24 jaar), effecten van verandering in de opname praktijk (bijvoorbeeld een dummy variabele die het effect van de Diefstal Act in 1968 oppikt) en het aantal politiemensen. In dit hoofdstuk tonen we resultaten die sterke effecten van de "business cycle" op diefstal en diefstal met inbraak laten zien. We vinden geen aparte cycli in de criminaliteit reeksen. De cyclische bewegingen in criminaliteit kunnen volledig worden verklaard door de effecten van de "business cycle". De verklarende variabelen hebben meer effect op de korte termijn dan op de lange termijn van de dynamiek in criminaliteit ontwikkelingen over de tijd.

De derde contributie van dit hoofdstuk is onze onderzoeksaanpak waarin motivatie en opportuniteitseffecten (vanuit economisch perspectief) van diefstal en diefstal met inbraak worden onderscheiden. De criminele motivatie werkt door op het effect van de wisselende economische voorwaarden aan sociale spanning en sociale controle. Daarnaast werkt de criminele opportuniteit door in de beschikbaarheid en kwetsbaarheid van criminele doelen op basis theorie van routine activiteiten. De twee effecten zijn tegenstrijdig met elkaar, lopen door elkaar heen en vinden niet noodzakelijkerwijs op hetzelfde tijdstip plaats.

Criminaliteit is wijdverbreid in de groep van jonge personen, vooral jonge mannen. In een longitudinaal onderzoek van 411 mannen uit Londen (welke zijn gevolgd vanaf hun zevende levensjaar tot en met volwassenheid), vinden Farrington et al. (1998) dat een gemiddelde criminele carrière op een leeftijd tussen de 14 en 21 jaar begint, tien jaar duurt en rond de leeftijd van 26 jaar eindigt. Deze relatie tussen leeftijd en criminaliteit is robuust over de tijd, tussen landen, binnen demografische groepen en voor verschillende soorten van criminaliteit en het hangt niet af van de wijze waarop criminaliteit wordt gemeten (zie Hirschi et al., 1983). In de VS worden vergelijkbare resultaten gevonden. Grogger (1998) rapporteert dat 35 percent van alle mannen uit Philadelphia die in 1945 waren geboren, vóór hun 18de levensjaar werden gearresteerd, en dat een derde van alle mannen uit Californië die in 1956 geboren waren, tussen de leeftijd van 18 en 30 jaar zijn gearresteerd. Volgens het bureau voor de statistiek (Census) van de VS zaten in 1990 1.1 miljoen mensen in de gevangenis, en veel van deze mensen waren mannen tussen de 20 en 30 jaar oud.

In hoofdstuk 4 stellen wij dat vroege school-verlating (en het hieruit voortvloeiende gering aantal jaren van onderwijs) één van de belangrijkste oorzaken van hoge criminaliteit bij jonge mensen is. Dit heeft tot gevolg dat deze jongeren een slechte positie op de arbeidsmarkt verwerven en een afweging gaan maken tussen legale en illegale activiteiten om aan inkomsten te komen. Methodologisch is het lastig om een effect van het onderwijs op de criminaliteit te isoleren. Het is moeilijk om empirisch de causaliteit aan te tonen van onderwijs op criminaliteit (en niet andersom). Om dit probleem te behandelen adopteren wij een quasi-experimentele aanpak waarin we de variatie over de tijd van de leerplicht leeftijd in Engeland en Wales gebruiken om de richting van de causaliteit te bevestigen. We bestuderen de relatie tussen criminaliteit en onderwijs met het gebruik van twee Britse data bronnen. Het eerste is de "Offenders Index Database" (OID) dat alle misdaadplegingen in Engeland en Wales registreert en dat we koppelen aan de "Labour Force Survey" (LFS) met data over het aantal jaren onderwijs in bepaalde leeftijdsgroepen gemeten over de tijd. De tweede bron van gegevens bestaat uit de vrijheidsstraf cijfers uit de Census 2001, waarbij we kijken naar de criminaliteit en het genoten onderwijs in een grote cross-sectie van de Britse bevolking.

Onze resultaten laten zien dat er grote effecten bestaan van het genoten onderwijs op het criminele gedrag. De empirische resultaten zijn gebaseerd op de instrumentele variabelen methode. De resultaten wijken niet veel af van resultaten gebaseerd op de kleinste kwadraten methode ook al kunnen we geen causale interpretatie koppelen aan de laatstgenoemde methode. In onze empirische modellen van vermogensdelicten rapporteren wij het significante resultaat dat een laag niveau van onderwijs (of geen onderwijs) tot een hoog niveau van criminaliteit kan leiden. Wij bevestigen deze resultaten met de schattingen uit de cross-sectie van de 2001 Census. De implicaties van onze resultaten zijn duidelijk: ze laten zien dat de verhoging van menselijk kapitaal accumulatie verkregen door het langer volgen van onderwijs effectief is om de individuele positie op de arbeidsmarkt te verbeteren. Hieruit volgt dat de verhoging van de menselijk kapitaal accumulatie door jongeren langer onderwijs te laten volgen een speerpunt van beleid zou moeten zijn bij het bestrijden van de criminaliteit. In dit verband schrijft Eide (2000) dat "young people are perhaps not different, just poorer."

In hoofdstuk 5 onderzoeken wij de negatieve samenhang tussen onderwijs en criminaliteit door gebruik te maken van gegevens van een steekproef van jonge Australische tweelingen. Leidt crimineel gedrag tot een lager onderwijsniveau of leidt meer onderwijs juist tot minder criminaliteit? Wij gebruiken twee aspecten van de Australische gegevensbronnen over onderwijs en criminaliteit om deze vraag te beantwoorden. Ten eerste kunnen wij, omdat het gegevens van tweelingen betreft, controleren voor vele niet waargenomen karakteristieken die zowel van invloed zijn op crimineel gedrag als op school carrière. Ten tweede kunnen wij, aangezien crimineel gedrag op verschillende momenten in de tijd gemeten is – zowel voor als na de middelbare school – de causaliteit tussen criminaliteit en onderwijs onderzoeken. Omdat vroeg-crimineel gedrag de vorming van menselijk kapitaal kan beïnvloeden, en menselijk kapitaal van invloed kan zijn op crimineel gedrag later in het leven, gebruiken wij een twee-stapsanalyse.

In de eerste stap onderzoeken wij de relatie tussen vroege criminaliteit en de accumulatie van menselijk kapitaal. Wij schatten het effect van vroege arrestaties (vóór het 18de levensjaar) op het aantal jaren genoten onderwijs met gebruikmaking van de binnen-tweelingparen ("within-twin") schattingsmethode, waarbij er rekening wordt gehouden met genetische en sociaal-economische factoren die gelijk zijn voor beide tweelingen van het tweelingpaar. Verder kijken wij of het tijdstip van de arrestaties belangrijk is voor het bereikte onderwijsniveau. In de tweede stap schatten wij het effect van het bereikte onderwijsniveau op drie indicatoren van criminaliteit: vrijheidsstraf, aantal arrestaties vanaf de leeftijd van 18 jaar en het totaal aantal arrestaties. Vroeg-crimineel gedrag kan hoog gecorreleerd zijn met andere verklarende variabelen en zo een verstorende invloed hebben op de schattingsmethode. Daarom controleren wij voor het aantal vroege arrestaties met een maat voor gedragsstoornissen binnen tweelingparen.

Wij concluderen dat arrestaties op jonge leeftijd (vóór het 18de levensjaar) een sterk negatief effect hebben op het bereikte onderwijsniveau, ook binnen tweelingparen. Arrestaties op de leeftijd van 13, 14 of 15 jaar hebben het meest negatieve effect op het onderwijsniveau. Wij vinden ook dat onderwijs criminaliteit reduceert. De invloed van onderwijs op criminaliteit is voor circa driekwart toe te schrijven aan het feit dat arrestaties op jonge leeftijd en een antisociale gedragsstoornis de deelname aan onderwijs verminderen. Dit betekent dat de sterke samenhang tussen onderwijs en criminaliteit grotendeels valt toe te schrijven aan het effect van vroeg-crimineel gedrag op onderwijs. Vroeg-crimineel gedrag heeft ook een sterk effect op crimineel gedrag op latere leeftijd. Programma's gericht op de preventie van vroeg-crimineel gedrag kunnen dan ook hoge sociale en privé baten opleveren.

Dit hoofdstuk bevat verschillende bijdragen aan de economische literatuur over de relatie tussen onderwijs en criminaliteit. Ten eerste gebruiken wij de longitudinale karakteristieken van onze data om een relatie tussen menselijk kapitaal en criminaliteit in beide richtingen te schatten. Deze identificatie strategie is niet eerder in deze context gebruikt. Ten tweede onderzoeken wij het effect van vroeg-crimineel gedrag op menselijk kapitaal accumulatie terwijl wij controleren voor de fixed effects (FE) binnen de tweelingparen. Wij kennen geen eerdere studies in de economische literatuur die het causale effect van vroeg-crimineel gedrag op onderwijs op deze wijze proberen te schatten. Ten derde bestaat er in de economische literatuur een groeiende interesse naar de effecten van vroege condities in het leven op volwassenheidskarakteristieken (Currie and Stabile 2006, 2007; Borghans et al., 2008). Dit hoofdstuk onderzoekt dezelfde vraagstukken.

In hoofdstuk 6 onderzoeken wij de langetermijneffecten van een antisociale gedragsstoornis op jonge leeftijd. Gekeken is naar het effect op het bereikte onderwijsniveau en op agressief en crimineel gedrag op latere leeftijd. Ook in dit hoofdstuk is gebruik gemaakt van gegevens van Australische tweelingen. Wij definiëren agressief en crimineel gedrag als negatief menselijk kapitaal ("building capital in anti-social or criminal activities," Currie and Stabile, 2007). De literatuur laat zien dat veel kinderen gedragsstoornissen en geestelijke gezondheidsproblemen hebben die nadelig zijn voor hun normale ontwikkeling en functioneren. Psychische angsten, gedragsstoornissen, concentratieproblemen en hyperactiviteit (ADHD) en depressie zijn de meest voorkomende stoornissen. Er wordt bijvoorbeeld geschat dat 500,000 (14%) van de Australische kinderen tussen de 4 and 18 jaar significante geestelijke gezondheidsproblemen hebben (Australian Centre for Community Child Health, 2006); dat ongeveer 50% 'moeilijk te behandelen' voorschoolse leeftijd kinderen een grote kans op verdere problemen in het leven hebben en daarvoor vaak professionele hulp inroepen; dat ongeveer een op de vijf kinderen en adolescenten in de VS geestelijke gezondheidsproblemen kunnen hebben (Currie and Stabile, 2006). Ondanks deze hoge cijfers is er relatief weinig bekend over de effecten van deze geestelijke gezondheidsproblemen bij kinderen op de langere termijn.

Als iemand gedrag vertoont dat overeenkomt met "a repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate societal norms or rules are violated" (APA, 1994), dan valt hij/zij in de categorie van gedragsproblemen. Een gedragsstoornis staat bekend als een 'verstorende gedragsstoornis' omdat deze een effect heeft op kinderen, hun families, buren en schoolgenoten, en meestal gepaard gaat met latere delinquentie en criminele activiteiten. De antisociale gedragsstoornis is gemeten met een indicatorenlijst die is gebaseerd op diagnostische criteria uit de psychiatrie (zoals agressie tegen mensen of dieren, vernielen van eigendom, diefstal of bedrog en ernstige overtreding van regels). Net als Currie and Stabile (2007) schatten wij de effecten van de antisociale gedragsstoornis op het positief menselijk kapitaal (blijven zitten, cijfers op de lagere en middelbare school, totaal aantal jaren van onderwijs) en het negatief menselijk kapitaal (aantal arrestaties, vrijheidsstraffen en fysieke aanvallen op andere mensen).

De twee belangrijkste problemen bij het schatten van de lange termijneffecten van gedragsstoornissen bij kinderen op het bereikte onderwijsniveau en op agressief en crimineel gedrag zijn het meten van de gedragsstoornis en niet waargenomen kenmerken ("omitted variable") bias. Dit hoofdstuk levert twee bijdragen aan het oplossen van deze problemen. Ten eerste is het eenvoudig na te gaan of een kind koorts heeft, maar gedragsstoornissen van kinderen zijn moeilijk te identificeren. Wij gebruiken hiervoor de zelf-rapportages van de volwassen tweelingen op 21 vragen, die de criteria voor een gedragsstoornis volgens de definitie van de American Psychiatric Association (APA, 1994) volgen. Een andere bijdrage van dit hoofdstuk is dat we informatie hebben over de leeftijd waarop de gedragsstoornis voor het eerst plaatsvond. Wij gebruiken deze informatie om het effect van het tijdstip van de gedragsstoornis op menselijk kapitaal te onderzoeken. Ten tweede kunnen de schattingen van het effect van de gedragsstoornis op de menselijk kapitaal accumulatie door de niet waargenomen kenmerken van de kinderen en hun families beïnvloed zijn. In dit hoofdstuk schatten wij de binnen-tweelingparen modellen met gebruikmaking van de gegevens over de (identieke) tweelingen. Het voordeel van het gebruik van tweelingen data ten opzichte van gegevens van 'gewone' broers en zussen is dat de familie omstandigheden van de tweelingen meer gelijk zijn dan de familie omstandigheden van 'gewone' broers en zussen. Belangrijker is het feit dat de identieke tweelingen genetisch identiek zijn, terwijl 'gewone' broers en zussen, gemiddeld, slechts ongeveer de helft van de genetische factoren delen. Het gebruik van de data over de (identieke) tweelingen kan vertekening door niet geobserveerde factoren dus reduceren.

Het hebben van een antisociale gedragsstoornis blijkt een sterk effect te hebben op zowel het verwerven van menselijk kapitaal als op agressief en crimineel gedrag gedurende het leven. Een antisociale gedragsstoornis vermindert de kans op het halen van een diploma van het voortgezet onderwijs (high school) binnen paren van eeneiige tweelingen van het landelijke gemiddelde van 79 procent naar ongeveer 66 tot 75 procent, en verhoogt de kans om ooit gearresteerd te worden van ongeveer 2 procent naar 9 tot 18 procent. Gevoeligheidsanalyses suggereren dat de effecten mogelijk nog sterker zijn. Ook is gevonden dat de effecten sterker zijn naarmate de antisociale gedragsstoornis zich manifesteert op jongere leeftijd. Wij concluderen dat geestelijke gezondheidsproblemen bij jongeren grote menselijke en financiële kosten voor families en de maatschappij met zich meebrengen. Effectieve behandelingen op vroege leeftijd kunnen dan ook grote opbrengsten opleveren.

Samenvattend kunnen we de volgende conclusies trekken uit dit proefschrift. Uit een macro-economische analyse blijkt dat strenger straffen weliswaar de totale criminaliteit reduceert maar dat het effect niet zo groot is als andere beleidsmaatregelen om de criminaliteit te reduceren. Ten tweede reageert de totale criminaliteit op economische "business cycle" condities. We hebben laten zien dat de trends in de criminaliteit goed door economische factoren verklaard kan worden. Ten derde hebben we in een micro-economische analyse aangetoond dat er een sterke samenhang bestaat tussen onderwijs en criminaliteit. Het blijkt dat meer onderwijs het niveau van de criminaliteit onder jonge mensen reduceert. Als onderwijs daadwerkelijk criminaliteit reduceert, dan is de sociale opbrengst hoger dan de persoonlijke opbrengst door onderwijs. De sterke relatie tussen onderwijs en criminaliteit wordt beïnvloed door het effect van vroege criminaliteit op het bereikte onderwijsniveau. Crimineel gedrag op jonge leeftijd leidt tot een lager onderwijsniveau, en het lagere onderwijsniveau leidt weer tot meer crimineel gedrag. Ten vierde zijn jongeren die een antisociale gedragsstoornis hebben veel minder succesvol in het onderwijs dan hun leeftijdsgenoten. Ook blijken zij op latere leeftijd vaker agressief of crimineel gedrag te vertonen. Wij concluderen dat geestelijke gezondheidsproblemen tijdens de jeugd tot grote menselijke en financiële kosten voor families en de maatschappij kunnen leiden. Programma's die probleemgedrag op jonge leeftijd weten te verminderen kunnen daarom hoge opbrengsten opleveren, zowel voor het individu als voor de samenleving.

Sažetak (Summary in Serbian)

Od jednih do drugih izbora, individualni kriminal je jedan od najvećih socijalnih i političkih problema, kao i jedna od omiljenih tema na listi obećanja političara u toku izborne trke. Individualni kriminal nameće društvu visoke troškove, koji obično ne ulaze u zvanične pokazatelje blagostanja i nivoa sreće različitih zemalja. Polazeći od seminalnih radova Garija Bekera (Becker, 1968) i Isaka Ehrliha (Ehrlich, 1973), činjenica da se individualni kriminal može objasniti ekonomskim uslovima i motivima je široko prihvaćena u ekonomiji. U skladu sa standardnim ekonomskim modelom individualnog kriminala, odluka da se pojedinac bavi kriminalom zavisi od sveukupnog prihoda od legalnih aktivnosti, sklonosti ili 'ukusa' pojedinca prema kriminalu, kao i egzogenih faktora koji utiču na verovatnoću hapšenja i primenjene kaznene mere. Uprkos brojnih radova iz oblasti ekonomije i kriminologije na ovu temu, čini se da javnost uglavnom nije u toku sa ekonomskim pristupom problemu individualnog kriminala i njegovih posledica.

Sa ekonomske tačke gledišta, individualni kriminal se može lako objasniti modelom ponude i tražnje, u kome, s jedne strane, prekršilac zakona nudi određeni nivo kriminala, s druge strane, društvo traži zaštitu od kriminala, a država snabdeva traženu javnu zaštitu. Ponuda kriminala je rezultat izbora između legalnih i ilegalnih aktivnosti. Izbor pojedinca zavisi od neto prihoda od ilegalnih aktivnosti, koji je dobitak nakon što se pokriju svi troškovi neophodni za izvršenje pojedinog krivičnog dela. Prihod od kriminalne aktivnosti zavisi od tipa krivičnog dela i sklonosti ili 'ukusa' pojedinca prema kriminalu. Prihod od kriminalne aktivnosti može biti finansijske ili psihičke prirode (uzbuđenje zbog opasnosti, prihvatanje od strane prijatelja i istomišljenika, osećaj dostignuća ili ispunjenje seksualnog nagona u slučaju silovanja). Troškovi vezani za izvršenje krivičnog dela mogu biti direktni i indirektni (oportunitetni) troškovi. Direktni troškovi kriminala su materijalni troškovi (oprema, oružje, prevozna sredstva), psihički troškovi (griža savesti, nespokojstvo, strahovanje, nesklonost ili rizik) i očekivani troškovi primenjene kaznene mere (novčane kazne, odlazak u zatvor i ostale kaznene mere, kao i novčani troškovi vezani za parnice i odlazak na sud). Oportunitetni troškovi kriminala su neto prihodi (prihodi minus troškovi) od propuštenih legalnih aktivnosti tokom planiranja, izvršenja i prikrivanja krivičnog dela. Na primer, obrazovanje i socijalna pomoć utičnu na rast oportunitetnih troškova kriminala, zato što uvećavaju potencijalne prihode od legalnih aktivnosti. Ponuda kriminala je pozitivno korelisana sa neto prihodima od kriminalne aktivnosti, što znači da će (potencijalni) izvršilac krivičnog dela učestati svoje aktivnosti kada neto prihodi od kriminala rastu. Primeri uslova koji mogu da uzrokuju da se kriva

ponude kriminala rotira u desno su demografski uslovi (na primer, uvećana proporcija omladine u društvu), nedostatak legalnih mogućnosti za zaposlenje ili olakšavanje krivičnih mera za pojedina krivična dela. Dok se imovinski kriminal najlakše objašnjava ekonomskim faktorima, različiti tipovi nasilnog kriminala, koji se dešavaju kao sporedni proizvod kriminalnog dela koji je materijalno motivisan, se isto mogu objasniti ekonomski faktorima.

Ova doktorska disertacija se uglavnom fokusira na ponudu kriminala, gde se relacija između kriminala i različitih ekonomskih i socijalnih faktora analizira primenom ekonometrijskih modela na agregatne (nacionalne) vremenske serije (podaci za Virdžiniju u Sjedinjenim Američkim Državama i Englesku i Vels u Velikoj Britaniji), agregatne uporedne podatke (podaci za Englesku i Vels u Velikoj Britaniji) i individualne anketne podatke (podaci za Australiju). Primarni cilj ovog rada je da empirijski prikaže relaciju između kriminala i različitih ekonomskih i socijalnih faktora. Zaključak ovog rada je da ako se individualni kriminal posmatra kao rezultat racionalnog ponašanja pojedinca, onda bi ekonomske i socijalne mere morale da budu glavne mere društvene zaštite i prevencije od kriminala.

Poglavlje 2 istražuje kretanje kriminala u Virdžiniji u Sjedinjenim Američkim Državama (SAD). Gledano kroz istoriju, trend kriminala u SAD je bio u porastu tokom 1960-ih i 1970-ih, da bi se u 1980-im radikalno promenio, tako da u 1990-im godinama vidimo dramatični pad kriminala u svim kategorijama. Levit (Levitt, 2005) piše da su 1960-e i 1970-e bile "vrhunac liberalnog pravnog sistema i poboljšanja prava počinilaca krivičnih dela;" verovatnoća kažnjavanja je bila toliko niska da jednostavno nije koštalo previše (u ekonomskom smislu) da se postane kriminalac. Međutim, tokom 1980-ih se ovaj trend radikalno menja. Prava počinioca krivičnih dela su postala ograničena i došlo je do pooštravanja kazni za izvršena krivična dela. Tokom 1990-ih su skoro sve države u SAD usvojile jedan ili više obavezujućih krivičnih zakona. Svi ovi zakoni su imali jedan imenitelj: pooštravanje obavezujućih zakonskih kazni odlaskom u zatvor za pojedince koji su već bili osuđeni na zatvorski pritvor. Iako se očekivalo da je pooštravanje obavezujućih krivičnih zakona uticalo na dramatični pad kriminala tokom 1990-ih, među kriminolozima i ekonomistima je debata u toku sve do danas u vezi faktora koji su tome doprineli. Najčešće se kao razlozi navode legalizacija abortusa dvadeset godina ranije, privredna ekspanzija, lokalne socijalne i ekonomske mere, promene na tržištu kreka i drugih opijata i/ili pooštravanje obavezujućih krivičnih zakona. Ovo poglavlje doprinosi boljem razumevanju relacije između anti–kriminalnih nastojanja od strane države i redukcije kriminala tokom vremena, primenom ekonometrijskih modela na podatke za Virdžiniju u SAD.

U ovom poglavlju posmatramo efekat nove zakonske krivične mere u Virdžiniji na registrovane podatke imovinskog kriminala (provalna krađa, krađa, krađa vozila i pljačka) i nasilnog kriminala (otežani napad, ubistvo i silovanje), korišćenjem različitih modela za analizu vremenskih serija. Virdžinija je ukinula uslovni otpust i reformisala krivični zakon za sve počinioce krivičnih dela počev od prvog januara 1995. godine. Rezultat ove nove zakonske mere je bilo značajano pooštravanje kazne za sve počinioce nasilnog kriminala (naročito za silovanje i ubistvo), kao i za one koji su u prošlosti počinili neko nasilno krivično delo. Iako se na počinioce imovinskog kriminala nije direktno uticalo novim zakonom, očekivalo se da se broj zatvorenih osoba poveća i u ovoj kategoriji kriminala (naročito za provalnu krađu i pljačku). Ako pretpostavimo da stepen oštrine krivičnog zakona služi kao mera zastrašivanja potencijalnim počiniocima krivičnih dela, očekujemo da su anti–kriminalna nastojanja od strane države Virdžinije negativno uticala na registrovane stope kriminala, naročito na stope nasilnog kriminala.

Ovo poglavlje ispituje efekat nove zakonske mere u Virdžiniji na registrovane stope kriminala korišćenjem različitih modela za analizu vremenskih serija. Nova zakonska mera se tretira kao pomak nivoa, gde se nivo posmatrane vremenske serije permanentno menja u trenutku kada je novi zakon stupio na snagu. To znači da se nova zakonska mera tretira kao *intervencija*, koja se u model uvodi putem veštačke varijable (jednaka nuli pre januara 1995. godine i jedinici posle januara 1995. godine.) Intervencioni efekat nove zakonske mere u Virdžiniji na registrovane stope kriminala se ocenjuje strukturnim modelom vremenskih serija (STS) (Harvey i Koopman, 1997) i autoregresionim integrisanim procesom pokretnih proseka (ARIMA) Boksa i Dženkinsa (Box i Jenkins, 1976). Oba pristupa predstavljaju uspešne metode za intervencionu analizu. Strukturni modeli vremenskih serija se sastoje od komponenti kao što su trend, sezona, ciklus i iregularna komponenta, koji su jednostavni za interpretaciju i omogućavaju ocenu vremenski zavisnih regresionih parametara. U okviru STS modela ocenjujemo jednostruki, kao i simultani strukturni model vremenskih serija, koji omogućava uvođenje kontrolne grupe u analizu.

Postojeća literatura nudi ograničenu podršku tezi da je nova zakonska mera u Virdžiniji služila kao mera zastrašivanja potencijalnim počiniocima krivičnih dela. Kesler i Levit (Kessler i Levitt, 1999) objašnjavaju da su prethodna istraživanja u ovoj oblasti, koja su imala za cilj analizu efekata pooštravanja krivičnih mera u SAD tokom 1990-ih godina, došla do konfliktnih zaključaka. Ovi radovi vrlo često pokazuju slab ili nepostojeći efekat novih zakonskih mera. Do sličnog zaključka dolazi i analiza u ovom poglavlju.

Ovo poglavlje doprinosi postojećoj literaturi u sledećem. Prvo, empirijska analiza predstavlja sveobuhvatnu evaluaciju nove zakonske mere u Virdžiniji. Zvanični izveštaji u Virdžiniji sve do danas ne uspevaju da dokažu da li su realizovane redukcije kriminala u toku 1990-ih bile uzrokovane postojećim anti–kriminalim inicijativama. Drugo, vidimo ovo poglavlje kao uvod i demonstraciju fleksibilnosti (jednostruki i simultani pristup) strukturnih modela vremenskih serija u oblasti analize kriminalnih podataka. Do sada su vremenske serije iz oblasti kriminala uglavnom bile analizirane uz pomoć ARIMA procesa. Treće, interventiona analiza efekta nove zakonske mere u Virdžiniji pokazuje superiornost STS i ARIMA modela u odnosu na jednostavne regresione modele.

Poslovni ciklus ima prožimajući efekat na strukturu ekonomskih prilika i samim tim na ponašanje pojedinca. Ovaj efekat se odražava na različite socijalne indikatore, kao na primer broj đaka u školama, stopu rađenja ili stopu učešća radne snage. Polazeći od ovakvog stanovišta, možemo očekivati da poslovni ciklus utiče i na kretanje stope kriminala u društvu. Postojeća literatura pokazuje da se imovinski kriminal kontra-ciklično ponaša, opadajući i rastući zajedno sa ekonomskim bumom i padom. S druge strane, poslovni ciklus ne utiče značajno na nasilni kriminal, sem na tipove nasilnog kriminala koji imaju očiglednu finansijsku motivaciju (kao na primer, pljačka).

Poglavlje 3 ocenjuje efekat poslovnog ciklusa na stopu imovinskog kriminala (provalna krađa i krađa) u Engleskoj i Velsu u Velikoj Britaniji tokom druge polovine dvadesetog veka. U Velikoj Britaniji je stopa kriminala takođe rasla tokom 1970-ih i 1980-ih godina, tako da je početkom 1990-ih dostigla rekordan nivo i postala značajan društveni problem. Porast stope kriminala je bio znatno veći nego u ostalom delu Evrope. Osim toga, stopa imovinskog kriminala u Velikoj Britaniji je rasla u toku 1970-ih, 1980-ih i ranih 1990-ih godina. Postojeća literatura pokazuje da je stopa imovinskog kriminala u Velikoj Britaniji tokom 1990-ih bila veća nego u SAD u istom periodu (Machin i Meghir, 2004).

U cilju analize efekata poslovnog ciklusa na kretanje registrovane stope provalne krađe i krađe u Engleskoj i Velsu u periodu od 1955. do 2001. godine, koristimo jednostruki i simultani strukturni model vremenskih serija. Poslovni ciklus se tretira kao stacionarni autoregresivni proces pokretnih proseka (ARMA), što je prvi doprinos našeg pristupa problemu efekta poslovnog ciklusa na registrovane stope kriminala. Simultani strukturni model vremenskih serija istovremeno analizira vremenske serije kriminala (stopa provalne krađe i krađe), zajedno sa ekonomskim (stopa bruto društvenog proizvoda i nezaposlenosti) i varijablama zaštite i zastrašivanja (broj policijskog kadra). Simultana analiza kriminalnih, ekonomskih i varijabli zaštite i zastrašivanja predstavlja drugi doprinos tekućeg poglavlja postojećoj literaturi. Kao eksplanatorne varijable u model ulaze varijable krivično-zakonodavnog sistema (kaznena stopa, dužina zatvorske presude i stopa zatvaranja), demografske varijable (muški pol, omladina uzrasta 15-24 godina), varijable efekta promene u načinju registrovanja kriminala (veštačka varijabla koja uzima u obzir uvođenje Dekreta o Krađi u 1968-oj godini) i varijable zaštite i zastrašivanja (broj policijskog kadra). Analiza pokazuje signifikantnu zavisnost kretanja stopa provalne krađe i krađe od kretanja poslovnog ciklusa. S druge strane, ne nalazimo podršku pretpostavci da postoji poseban, kriminalni ciklus. Dalje, objašnjavajuće varijable značajnije utiču na kratkoročno nego na dugoročno kretanje analiziranih stopa kriminala.

Treći doprinos ovog poglavlja postojećoj literaturi je da model otkriva motivacione i oportunističke motive za izvršenje imovinskog krivičnog dela. Kriminalna motivacija operiše kroz efekat promenljivih ekonomskih uslova na socijalni pritisak i socijalnu kontrolu (teorije socijalnog pritiska i socijalne kontrole), dok kriminalni oportunizam utiče na raspoloživost i ranjivost kriminalnih ciljeva kroz teoriju rutinske delatnosti. Ova dva efekta deluju u suprotnom smeru i ne moraju da se istovremeno realizuju.

Kriminal je široko rasprostranjen među omladinom, naročito među mlađim muškarcima. Analiza Faringtona i drugih (Farrington i drugi, 1998), koja prati 411 muškaraca u centralnom Londonu od njihove sedme godine do odrastanja, pokazuje da prosečna kriminalna karijera počinje negde između 14-e i 21-e godine starosti, traje deset godina, završava se oko 26-e godine i (za one koji su počinili više od jednog krivičnog dela) sadrži u proseku 4.6 prekšaja zakona koji dovode do krivične presude i kazne. Ova relacija između uzrasta i kriminala je stabilna tokom vremena, između različitih zemalja, demografskih podgrupa, tipova kriminalnog dela i ne zavisi od načina na koji se kriminal meri (Hirschi i drugi, 1983). Za SAD važe slična zapažanja. Groger (Grogger, 1998) iznosi da tridesetpet procenata svih muškaraca rođenih 1945. godine u Filadelfiji je bilo uhapšeno pre svoje 18-e godine, dok je jedna trećina svih muškaraca rođenih 1956. godine u Kaliforniji bila uhapšeno između svoje 18-e i 30-e godine starosti. Popis u SAD pokazuje da je u 1990. godini 1.1 milion (0.4%) ljudi bilo u zatvoru, od čega su to mahom bili muškarci u svojim dvadesetim i tridesetim godinama.

U poglavlju 4 diskutujemo o tome kako je jedan od glavnih uzroka velikog broja kriminalnih dela među mladim ljudima uslovljen ranim napuštanjem škole i smanjenim brojem godina obrazovanja, što se onda odražava na individualnu radnu produktivnost i mogućnost sticanja dohotka na legalan način. Konceptualno, postoje poteškoće u izolovanju uticaja obrazovanja na kriminalnu aktivnost. Posebno, teško je osigurati da pravac uzročnosti ide od obrazovanja ka kriminalu (a ne u suprotnom smeru). Da bi dokazali pravac uzročnosti, usvajamo kvazi-eksperimentalni prisitup koji se oslanja na varijaciju u stečenom obrazovanju tokom vremena, prouzrokovanu promenama u zakonu o uzrastu kada se završava obavezno školovanje u Engleskoj i Velsu u Velikoj Britaniji. Relaciju između kriminala i obrazovanja analiziramo koristeći dva britanska izvora podataka. Prvi izvor predstavlja baza popisanih prestupnika (engl. Offenders Index Database ili OID) koja pokriva sve prestupnike u Engleskoj i Velsu, i nju uparujemo sa podacima istraživanja ankete o radnoj snazi (engl. Labour Force Survey ili LFS) po starosnoj dobi tokom vremena. Drugi izvor su podaci o broju zatvorenika iz Popisa za 2001. godinu, gde analiziramo relaciju između kriminala i obrazovanja na velikom uporednom preseku britanskog stanovništva.

Naši rezultati ukazuju na značajan uticaj obrazovanja na kriminal. Korišćenjem metoda instrumentalnih varijabli pokazujemo da su empirijske ocene, kada koristimo promenljivu obrazovanja koja je najpodesnija za ovaj pristup, slične onima koje su bazirane na metodu običnih najmanjih kvadrata i koje ne moraju da imaju kauzalnu interpretaciju. U empirijskom modelu imovinskog kriminala pokazujemo da je nizak stepen obrazovanja, a naročito nedostatak obrazovnih kvalifikacija, signifikantno povezan sa višom stopom krivičnih prekršaja. Ove rezultate potkrepljujemo rezultatima analizirane relacije između broja zatvorenika i nedostatka obrazovanih kvalifikacija na bazi uporednih podataka iz Popisa za 2001. godinu. Posledice ovih zaključaka su jasne i pokazuju da stimulisanje akumulacije ljudskog kapitala kroz sistem obrazovanja ili na neki drugi način koji se može pokazati kao efektivan za poboljšanje individualne radne produktivnosti, bi trebalo da bude ključna društvena mera u pokušaju da se utiče na smanjenje kriminala. Po rečima Eidea (Eide, 2000), "mladi ljudi možda nisu različiti, već samo siromašniji".

U poglavlju 5 se bavimo relacijom između ljudskog kapitala i kriminala na bazi uzorka mlađe kohorte australijskih blizanaca. Koristimo dva aspekta podataka australijskog anketnog istraživanja o vezi između obrazovanja i kriminala. Prvo, s obzirom da su podaci dobijeni od strane blizanaca, u stanju smo da kontrolišemo veliki broj neposmatranih (engl. unobserved) karakteristika koje utiču kako na kriminalno ponašanje pojedinca tako i na njegove/njene odluke vezane za obrazovanje. Drugo, pošto je kriminalno ponašanje mereno u različitim perioda, pre i nakon završetka srednje škole, možemo takođe ukazati i na uzročnu vezu između kriminala i obrazovanja. S obzirom da kriminalno ponašanje u ranom uzrastu može uticati na formiranje ljudskog kapitala, a da ljudski kapital može da utiče na kriminalno ponašanje u kasnijim periodima života, pristupamo analizi u dva koraka.

U prvom koraku izučavamo vezu između ranog kriminala i akumulacije ljudskog kapitala. Konkretno, ocenjujemo efekat hapšenja u uzrastu mlađem od 18 godina na sticanje obrazovanja koristeći takozvani 'unutar-par-blizanaca' (engl. within-twin) metod. Dodatno istražujemo da li vreme hapšenja utiče na sticanje obrazovanja. U drugom koraku istražujemo kako stečeno obrazovanje utiče na kriminal, gde koristimo tri mere kriminalne aktivnosti: broj zatvaranja, broj hapšenja nakon uzrasta od 18 godina i ukupan broj hapšenja. Kako maloletničko delikventno ponašanje može da bude važna objašnjavajuća promenljiva u ocenjenoj vezi između obrazovanja i kriminala, broj hapšenja u uzrastu mlađem od 18 godina i mera neprimerenog ponašanja kod parova blizanaca služe kao egzogene varijable u ovoj drugoj relaciji.

Ovo poglavlje daje doprinos ekonomskoj literaturi koja se bavi odnosom između obrazovanja i kriminala na nekoliko načina. Pre svega, dopunjujemo ne tako obimnu literaturu o kauzalnoj vezi između obrazovanja i kriminala i koristimo identifikacionu strategiju koja se ranije nije primenjivala u ovom kontekstu. Konkretno, koristimo longitudinalnu prirodu naših podataka da bi ocenili relaciju između kriminala i ljudskog kapitala u oba pravca. Na drugom mestu, istražujemo uticaj ranog kriminala na ulaganja u ljudski kapital dok kontrolišemo fiksne efekte (engl. fixed effects) kod parova blizanaca. Nismo upućeni u to da postoje istraživanja u ekonomskoj literaturi koja se, na ovaj način, bave kauzalnim efektom ranog kriminala na obrazovanje. Treće, postoji rastuće interesovanje u ekonomskoj literaturi o uticaju okolnosti u ranom dobu razvoja ličnosti na izgradnju odrasle osobe (Currie i Stabile, 2006, 2007; Borghans i drugi, 2008). Ovo poglavlje se bavi sličnim temama.

Zaključujemo da broj hapšenja u uzrastu mlađem od 18 godina ima statistički signifikantan efekat na akumulaciju ljudskog kapitala. Ovaj efekat je u najvećoj meri determinisan uzrastom kada je do hapšenja došlo, pokazujući da su hapšenja u dobi od 13, 14 ili 15 godina najštetnija za sticanje obrazovanja. Takođe zaključujemo da akumulacija ljudskog kapitala negativno utiče na kriminal. Međutim, kada u regresiji kontrolišimo broj hapšenja i meru neprimerenog ponašanja u ranom periodu odrastanja, ocenjeni uticaj ljudskog kapitala na kriminal se redukuje na manje od četvrtine prethodne ocene. Izuzetno štetni efekti ranog kriminala postaju očigledniji ako uzmemo u razmatranje ocenjene efekte ranog hapšenja (u uzrastu mlađem od 18 godina) na sve tri korišćene mere kriminala. U ovoj regresiji takođe ocenjujemo statistički signifikantan efekat ranog kriminalnog ponašanja na kasnije učestvovanje u kriminalu. Ovaj efekat je mnogo veći od (izolovanog) uticaja ljudskog kapitala na kriminal. Stoga, u ovom poglavlju, zaključujemo da je jaka veza između ljudskog kapitala i kriminala uglavnom determinisana efektom ranog kriminalnog ponašanja na obrazovanje. Programi koji uspevaju da spreče rano kriminalno ponašanje mogu da imaju značajan privatni i društveni doprinos.

Izvlačeći dodatne koristi iz podataka o australijskim blizancima, u poglavlju 6 analiziramo uticaj mentalnih problema i problema u ponašanju u ranom periodu odrastanja na akumulaciju ljudskog kapitala, gde tretiramo kriminal kao 'negativan ljudski kapital' ("kapital koji se gradi u anti-društvenim ili kriminalnim aktivnostima", Currie i Stabile, 2007). Literatura pokazuje da veliki broj dece ima mentalne probleme koji ih ometaju u normalnom razvoju i funkcionisanju. Nemir, poremećaji u ponašanju, nedostatak pažnje/hiperaktivnost (ADHD) i depresija su najčešći primeri mentalnih problema kod dece. Literatura pokazuje da, na primer, 500,000 (14%) australijske dece u dobi između 4 i 18 godina ima značajne mentalne probleme (Australian Centre for Community Child Health, 2006); velika je verovatnoća da će do 50% predškolske dece 'sa kojima se teško izlazi na kraj' imati i dalje poteškoće koje će zahtevati profesionalnu pomoć; u proseku jedno od petoro dece i adolescenata u SAD može imati mentalne poremećaje (Currie and Stabile, 2006). Uprkos ovim zabrinjavajućim brojkama, malo se zna o dugoročnim posledicama mentalnih problema i problema u ponašanju kod dece.

Ovo poglavlje analizira dugoročne efekte poremećaja u ponašanju tokom detinjstva na akumulaciju ljudskog kapitala i kasnijeg nasilnog i kriminalnog ponašanja u životu, koristeći podatke o australijskim blizancima. Ako neko pokazuje "ponavljajući i konzistentan obrazac ponašanja u kome su osnovna prava drugih ili glavne društvene norme koje priliče određenom uzrastu ugrožene" (APA, 1994), onda on/ona pripada kategoriji pojedinaca sa problemom poremećaja u ponašanju. Poremećaj ponašanja snažno utiče na samu decu i njihove porodice, susede i školu i često je povezan sa delikventnom ili kriminalnom aktivnošću. Merimo poremećaj ponašanja koristeći dijagnostičke kriterijume Američke psihijatrijske asocijacije (APA, 1994). Slično kao kod Kuri i Stabilea (Currie i Stabile, 2007), ocenjujemo efekat poremećaja ponašanja na pozitivan ljudski kapital, uključujući merila kao što su ponavljanje razreda, ocene u osnovnoj i srednjoj školi i sticanje obrazovanja, i na 'negativan ljudski kapital', uključujući merila kao što su broj hapšenja, broj zatvaranja, fizičko napadanje drugih lica i sl.

Dva glavna izazova u ocenjivanju dugoročnog efekta poremećaja u ponašanju u detinjstvu na akumulaciju ljudskog kapitala su korišćena mera poremećaja u ponašanju i pristrasnost ocene u posmatranoj relaciji zbog izostavljanja relevantnih varijabli (engl. omitted variable bias). Ovo poglavlje daje doprinos ekonomskoj literaturi baveći se upravo ovim problemima. Prvo, lako je znati kada dete ima temperaturu, ali se dečiji mentalni problemi i problemi u ponašanju teže uočavaju. U tu svrhu, koristimo anketu australijskih blizanaca u kojoj se odrasli blizanci (u prosečnoj dobi od 39 godina) samoizveštavaju o svom ponašanju u mladosti. Anketa se u ovom delu sastoji od 21-og pitanja koja prate definiciju poremećaja u ponašanju vodeći se kriterijumima Američke psihijatrijske asocijacije. Dodatni doprinos ovog poglavlja je da posedujemo informaciju o uzrastu u kome su se poremećaji u ponašanju desili. Koristimo ovu informaciju za istraživanje efekta trenutka ispoljavanja problematičnog ponašanja na akumulaciju ljudskog kapitala. Drugo, ocena efekta poremećaja u ponašanju na akumulaciju ljudskog kapitala može biti ugrožena neposmatranim razlikama (engl. unobserved heterogeneity) između dece i njihovih porodica. U ovom poglavlju takođe ocenjujemo 'unutar-par-blizanaca' model koristeći podatke o (jednojajčanim) blizancima. Prednost korišćenja blizanaca umesto braće i sestara se ogleda u tome da su porodični uslovi obično sličniji kod blizanaca nego kod braće i sestara. Ono što je još važnije je da su jednojajčani blizanci genetski identični, dok braća i sestre, u proseku, dele samo jednu polovinu svog genetskog nasleđa. Stoga, koristeći podatke o blizancima, a pogotovo one o genetski identičnim blizancima, možemo da redukujemo pristrasnost ocene koja je uzrokovana neposmatranim karakteristikama unutar porodica.

Kroz rezultate ovog poglavlja uočavamo veliki broj štetnih efekata poremećaja u ponašanju na pozitivan i 'negativan ljudski kapital', čak i među parovima jednojajčanih blizanaca. Poremećaji u ponašanju povećavaju verovatnoću ponavljanja razreda i nezavršavanja srednje škole. Uz to, poremećaji u ponašanju mereni pre uzrasta od 18 godina imaju jak uticaj na nasilno i kriminalno ponašanje nakon uzrasta od 18 godina. Još jedno važno otkriće je da je uticaj poremećaja u ponašanju štetniji ako se ovakvo ponašanje javi u ranijem dobu ljudskog života. Različite provere osetljivosti (engl. robustness analysis) ocenjenih rezultata, kao što je analiza osetljivosti koja se bavi problemom greške merenja ili analiza efekta propagacije unutar para blizanaca, nagoveštavaju da su naše glavne ocene možda na nižem nivou u odnosu na pravi efekat poremećaja u ponašanju na akumulaciju ljudskog kapitala. Zaključujemo da mentalni problemi u detinjstvu nameću velike ljudske i finansijske troškove kako za porodicu tako i za društvo u celini. Rani intervencioni programi bi stoga mogli da imaju značajan doprinos.

Sumirajući, ova teza donosi sledece zaključke. Na makro novou, ocenjujemo da stroga kaznena politika dovodi do smanjenja agregatne stope kriminala, ali su pronađeni efekti zanemarljivi u odnosu na druge mere smanjenja kriminala. Drugo, agregatne stope kriminala reaguju na kretanje poslovnog ciklusa, pružajući tako jasno ekonomsko objašnjenje za posmatrane trendove u kriminalu. Treće, na mikro nivou, nalazimo da dodatne godine školovanja među mladim ljudima značajno smanjuju njihovu kriminalnu aktivnost. Ako obrazovanje smanjuje kriminal, onda školovanje ima društveni doprinos koji se često ne uzima u obzir od strane pojedinaca i društva u celini. U tom slučaju, društveni doprinos obrazovanja može da bude veći od privatnog doprinosa obrazovanja. Međutim, jaka veza između ljudskog kapitala i kriminala može biti vođena efektom ranog kriminalnog ponašanja na sticanje obrazovanja. Četvrto, nalazimo veoma štetne efekte ranog poremećaja u ponašanju na akumulaciju pozitivnog i 'negativnog ljudskog kapitala', gde kriminal vidimo kao "kapital koji se gradi u anti-društvenim ili kriminalnim aktivnostima", (Currie and Stabile, 2007). Zaključujemo da mentalni i poremećaji u ponašanju u detinjstvu nameću velike ljudske i finansijske troškove kako za porodicu tako i za društvo u celini. Rani intervencioni programi bi stoga mogli da imaju značajan doprinos.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

404. H. SUNARTO, *Understanding the role of bank relationships, relationship marketing, and organizational learning in the performance of people's credit bank.*
405. M.Â. DOS REIS PORTELA, *Four essays on education, growth and labour economics.*
406. S.S. FICCO, *Essays on imperfect information-processing in economics.*
407. P.J.P.M. VERSIJP, *Advances in the use of stochastic dominance in asset pricing.*
408. M.R. WILDENBEEST, *Consumer search and oligopolistic pricing: A theoretical and empirical inquiry.*
409. E. GUSTAFSSON-WRIGHT, *Baring the threads: Social capital, vulnerability and the well-being of children in Guatemala.*
410. S. YERGOU-WORKU, *Marriage markets and fertility in South Africa with comparisons to Britain and Sweden.*
411. J.F. SLIJKERMAN, *Financial stability in the EU.*
412. W.A. VAN DEN BERG, *Private equity acquisitions.*
413. Y. CHENG, *Selected topics on nonparametric conditional quantiles and risk theory.*
414. M. DE POOTER, *Modelling and forecasting stock return volatility and the term structure of interest rates.*
415. F. RAVAZZOLO, *Forecasting financial time series using model averaging.*
416. M.J.E. KABKI, *Transnationalism, local development and social security: the functioning of support networks in rural Ghana.*
417. M. POPLAWSKI RIBEIRO, *Fiscal policy under rules and restrictions.*
418. S.W. BISSESSUR, *Earnings, quality and earnings management: the role of accounting accruals.*
419. L. RATNOVSKI, *A Random Walk Down the Lombard Street: Essays on Banking.*
420. R.P. NICOLAI, *Maintenance models for systems subject to measurable deterioration.*
421. R.K. ANDADARI, *Local clusters in global value chains, a case study of wood furniture clusters in Central Java (Indonesia).*

422. V. KARTSEVA, *Designing Controls for Network Organizations: A Value-Based Approach.*

423. J. ARTS, *Essays on New Product Adoption and Diffusion.*

424. A. BABUS, *Essays on Networks: Theory and Applications.*

425. M. VAN DER VOORT, *Modelling Credit Derivatives.*

426. G. GARITA, *Financial Market Liberalization and Economic Growth.*

427. E. BEKKERS, *Essays on Firm Heterogeneity and Quality in International Trade.*

428. H. LEAHU, *Measure-Valued Differentiation for Finite Products of Measures: Theory and Applications.*

429. G. BALTUSSEN, *New Insights into Behavioural Finance.*

430. W. VERMEULEN, *Essays on Housing Supply, Land Use Regulation and Regional Labour Markets.*

431. I.S. BUHAI, *Essays on Labour Markets: Worker-Firm Dynamics, Occupational Segregation and Workplace Conditions.*

432. C. ZHOU, *On Extreme Value Statistics.*

433. M. VAN DER WEL, *Risk-free Rate Dynamics: Information, Trading, and State Space Modelling.*

434. S.M.W. PHLIPPEN, *Come Close and Co-Create: Proximities in pharmaceutical innovation networks.*

435. A.V.P.B. MONTEIRO, *The Dynamics of Corporate Credit Risk: An Intensity-based Econometric Analysis.*

436. S.T. TRAUTMANN, *Uncertainty in Individual and Social Decisions: Theory and Experiments.*

437. R. LORD, *Efficient pricing algorithms for exotic derivatives.*

438. R.P. WOLTHOFF, *Essays on Simultaneous Search Equilibrium.*

439. Y.-Y. TSENG, *Valuation of travel time reliability in passenger transport.*

440. M.C. NON, *Essays on Consumer Search and Interlocking Directorates.*

441. M. DE HAAN, *Family Background and Children's Schooling Outcomes.*

442. T. ZAVADIL, *Dynamic Econometric Analysis of Insurance Markets with Imperfect Information.*

443. I.A. MAZZA, *Essays on endogenous economic policy.*

444. R. HAIJEMA, *Solving large structured Markov Decision Problems for perishable-inventory management and traffic control.*

445. A.S.K. WONG, *Derivatives in Dynamic Markets.*

446. R. SEGERS, *Advances in Monitoring the Economy.*

447. F.M. VIEIDER, *Social Influences on Individual Decision Making Processes.*

448. L. PAN, *Poverty, Risk and Insurance: Evidence from Ethiopia and Yemen.*

449. B. TIEBEN, *The concept of equilibrium in different economic traditions: A Historical Investigation.*

450. P. HEEMEIJER, *Expectation Formation in Dynamic Market Experiments.*

451. A.S. BOOIJ, *Essays on the Measurement Sensitivity of Risk Aversion and Causal Effects in Education.*

452. M.I. LÓPEZ YURDA, *Four Essays on Applied Microeconometrics.*

453. S. MEENTS, *The Influence of Sellers and the Intermediary on Buyers' Trust in C2C Electronic Marketplaces.*