FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

ILARIA MASIERO

THREE ESSAYS ON THE ECONOMICS OF CRIME

São Paulo
2017

FUNDAÇÃO GETULIO VARGAS

ESCOLA DE ECONOMIA DE SÃO PAULO

ILARIA MASIERO

THREE ESSAYS ON THE ECONOMICS OF CRIME

Tese submetida à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para a obtenção do título de Doutora em Economia.

Orientador: Prof. Dr. Rodrigo Reis Soares

São Paulo

2017

Masiero, Ilaria.
Three essays on the economics of crime / Ilaria Masiero. - 2017.
121 f.

Orientador: Rodrigo Reis Soares
Tese (doutorado) - Escola de Economia de São Paulo.

1. Crime - Aspectos sociológicos. 2. Criminologia. 3. Internet - Aspectos sociais. 4. Copa do mundo (futebol) - Brasil. 5. Polícia. 6. Segurança pública - Brasil. I. Soares, Rodrigo Reis. II. Tese (doutorado) - Escola de Economia de São Paulo. III. Título.

CDU 343.9(81)

ILARIA MASIERO

THREE ESSAYS ON THE ECONOMICS OF CRIME

Tese submetida à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para a obtenção do título de Doutora em Economia.

Data da aprovação: 19 de Maio de 2017.

Banca examinadora:

Prof. Dr. Rodrigo Reis Soares

(Columbia University e Escola de Economia de São Paulo)

Prof. Dr. Leandro Piquet Carneiro

(Universidade de São Paulo)

Prof. Dr. João Manoel Pinho de Mello

(Insper)

Prof. Dr. Vladimir Pinheiro Ponczek

(Escola de Economia de São Paulo)

Prof. Dr. André Portela Fernandes de Souza

(Escola de Economia de São Paulo)

To my unconventional family.

ACKNOWLEDGEMENTS

I am extremely grateful to my advisor, Rodrigo Soares, for his guidance throughout the elaboration of this thesis and for being an inspiring role model as a researcher and a professor.

I would also like to thank all the professors and fellow students at the São Paulo School of Economics who contributed to my understanding of economics. I am especially grateful for the comments and suggestions on my research I received from Paulo Arvate, Sammara Cavalcanti, Sérgio Firpo, Cristine Pinto, André Portela, Rafael Puglieri and many seminar and conference participants.

I thank the Public Safety Department of the State of São Paulo *(Secretaria de Estado da Segurança Pública do Governo de São Paulo*) for providing me with data about crime in the city of São Paulo, and Paulo Arvate for his precious help during the data-requesting process.

I gratefully acknowledge the financial support I received from the São Paulo School of Economics and from CAPES.

This whole chapter of my life would not have been possible without all the support and encouragement I have received from my loving family and friends, both in Italy and in Brazil. Above all, I thank Rafael for the love, adventure and companionship that have made these years exciting and smooth at the same time.

"The world is full of obvious things which nobody by any chance ever observes."

— *Arthur Conan Doyle, The Hound of the Baskervilles.*

ABSTRACT

Prevention is commonly recognized as the best way to go in the fight against delinquency. Yet, working out sound policies is an extremely challenging task. The first two essays in this thesis contribute to the pursuit of innovative tools by highlighting the crime-preventing potential of technologies and occupations that make people willingly alter their time allocation choices. The rationale behind this is simple: offenses will be averted if individuals choose to engage in less crime-conducive activities substituting time away from more crime-conducive activities. This mechanism is known in the literature as "voluntary incapacitation". In particular, the first paper analyzes the impact of Internet diffusion on crime in the US. Using a panel of state level yearly data and adopting an instrumental variable approach, I find a negative and significant relationship between Internet penetration and total and property crimes. Based on my theoretical framework, I interpret this outcome as reflecting voluntary incapacitation: time spent online crowds out alternative activities that would more likely lead to crime. The second essay investigates the entertainment-crime relationship by analyzing how criminal activity behaved in the city of São Paulo during the 2014 FIFA World Cup. Outcomes show that crime is significantly lower during matches, especially those with the highest remote viewership rates. Further tests suggest that these findings reflect the voluntary incapacitation effect, whereby the potential for criminal interaction (and thus crime) drops as people are busy watching the games. The main policy-relevant conclusion from the first two essays in this thesis is that providing access to technologies (such as the Internet) and entertainment activities (such as sporting events) may trigger a crime-preventive effect via voluntary incapacitation.

The third paper also relates to crime prevention by tackling a crucial issue in the economics of crime literature – the empirical assessment of the deterrent role of policing. The difficulty arises from the fact that crime and police presence are simultaneously determined, causing a problem of reverse causality. I address the issue by considering the natural experiment represented by the creation of a special police unit to intensify surveillance around a few tournament-related locations in the city of São Paulo during the 2014 FIFA World Cup. I take into account that the championship may impact crime in other ways than just through increased policing, namely via fan concentration and voluntary incapacitation. In order to disentangle the specific effect of police on crime, I exploit the fact that the World Cup affected different areas of the city through different channels and at different times.

Difference-in-differences estimates reveal that increased police presence leads to a significant reduction in criminal activity. The predicted elasticity of crime to police presence is remarkably close to estimates from previous studies.

Keywords: Crime • Voluntary incapacitation • Police • Internet • Entertainment • Time substitution effect • World Cup • Natural experiment.
JEL Classification: K42.

RESUMO

A prevenção é reconhecidamente o melhor caminho para o combate ao crime. Contudo, elaborar políticas preventivas é um grande desafio. Os primeiros dois ensaios dessa tese contribuem para a busca de ferramentas inovadoras neste âmbito ao explorar o papel de prevenção ao crime de tecnologias e formas de entretenimento que fazem as pessoas alterarem voluntariamente suas escolhas de alocação do tempo. O racional é simples: delitos podem ser evitados se as pessoas optarem por gastar seu tempo em atividades menos propícias à ocorrência de crimes tirando tempo de atividades mais propícias à ocorrência de crimes. Este mecanismo é conhecido na literatura como "*voluntary incapacitation*". O primeiro ensaio analiza o impacto da difusão da Internet sobre delitos nos EUA. Usando variáveis instrumentais em um contexto de painel por ano e estado, eu encontro uma relação negativa e significante entre penetração da Internet e crime total e crime contra a propriedade. Baseado no arcabouço teórico, eu interpreto este resultado como *voluntary incapacitation*: o tempo gasto *online* reduz o tempo gasto em atividades alternativas que mais provavelmente levariam a delitos. O segundo artigo investiga a ligação entre entretenimento e crime ao analizar a atividade criminosa na cidade de São Paulo durante a Copa do Mundo de 2014. Os resultados mostram que esta é significativamente menor durante os jogos, especialmente aqueles com as maiores taxas de audiência remota. Testes adicionais sugerem que estes resultados refletem o mecanismo de *voluntary incapacitation* uma vez que a possibilidade de interação entre potenciais vítimas e criminosos diminui enquanto as pessoas assistem os jogos. A maior contribuição para políticas públicas desses estudos é a conclusão que prover acesso a tecnologias (como a Internet) e atividades de entretenimento (como eventos esportivos) pode ajudar no combate ao crime através do efeito de *voluntary incapacitation*.

O terceiro ensaio também relaciona à prevenção ao crime e aborda um desafio da literatura, que é a quantificação empírica do efeito de dissuasão da polícia. A dificuldade nasce do fato de que presença policial e crime são determinados simultaneamente, causando um problema de endogeneidade. Eu abordo o assunto ao considerar o experimento natural representado pela criação de uma unidade especial de polícia para o monitoramento de algumas áreas específicas da cidade de São Paulo durante a Copa do Mundo de 2014. Eu levo em consideração que o campeonato deve influenciar o crime através de vários canais (concentração de torcedores e *voluntary incapacitation*), e não só por meio do aumento do efetivo policial. Para separar o efeito específico da polícia sobre crime, eu uso o fato de que a

Copa do Mundo impactou diferentes áreas da cidade em momentos diferentes e através de diferentes canais. Os resultados mostram que o aumento na presença da polícia leva a uma redução significativa da atividade criminosa. A elasticidade do crime à polícia estimada é muito próxima àquela calculada em outros estudos.

Palavras-chave: Crime • *Voluntary incapacitation* • Polícia • Internet • Entretenimento • Efeito de substituição de tempo • Copa do Mundo • Experimento natural.
Classificação JEL: K42.

Contents

List of Figures

List of Tables

I.

Internet diffusion: time online prevents crime?

Abstract

The impact of the Internet on crime is a priori ambiguous. I use a panel of state level yearly data on broadband (or high-speed Internet) penetration and criminal activity in the US to empirically estimate this relationship. I tackle the omitted variable problem by instrumenting high-speed Internet penetration with a measure of early-times telephone and cable TV diffusion, multiplied by year indicators. This approach originates from the consideration that two widespread broadband technologies use the pre-existing telephone and Cable TV networks to transmit data. These technologies accounted for almost the totality of connections at the beginning of the broadband era, and became less prevalent over time. Then, the pre-Internet penetration of telephone and cable TV is arguably correlated with broadband diffusion – and such correlation varies over time, providing a source of exogenous variation in broadband diffusion. Results show a negative and significant relationship between high-speed Internet penetration and total and property crimes. Relying on a theoretical framework derived from the results of previous literature, I interpret this outcome as reflecting a negative time substitution effect: time spent online crowds out alternative activities that would more likely lead to crime.

Keywords: Internet • Broadband • Crime • Time substitution effect.
JEL classification: K42.

I.1 Introduction

The Internet has revolutionized many aspects of society and people's lives, leaving room for social scientists to measure the depth and width of its aftermath. In line with this, scholars have put effort in studying the impact of the Web on a wide range of outcomes. This paper contributes to the literature by investigating one connection that has remained rather unexplored: the link between the Internet and crime.

The Web represents a virtual treasure trove of information, which is available at all times and with limited scope for access restriction. Exposure to such abundant supply of content, especially sensitive material such as violence or pornography, may affect criminal behavior in two opposite directions. On the one hand, there may be a positive arousal effect, which makes individuals more aggressive or prone to re-enact what they have seen online. On the other hand, a catharsis effect may be triggered, so that potential criminals substitute actual crime with online activities that inhibit or satisfy their impulses.

Although policy-makers at all levels – from heads of state to parents – seem to mostly worry about the content-driven impact of the Internet, there is another – possibly as important – mechanism through which the Web can affect criminal behavior, namely time substitution. Since the Internet became available, people are making room in their schedule for spending time online by substituting away from alternative occupations. This may impact criminal outcomes in either direction. It could be that the displaced activities trigger crime at a lower rate than time spent online – so that the Internet-crime relationship is positive – or it could be the other way around, so that spending time online reduces crime.

The mechanisms described above have been examined in previous empirical studies in psychology and economics. Overall, the literature suggests that, as far as the content channel is concerned, only the arousal effect is material – while there is very little evidence of a catharsis effect. Regarding the time substitution mechanism, previous research has overwhelmingly found its impact on crime to be negative and capable of offsetting the positive, content-driven arousal effect. These findings mainly arise from non-Internet-specific studies, which examine the impact on crime of sensitive content being transmitted over different technological platforms. In fact,

Internet-specific research is rather limited and non-conclusive, as it focuses on sex offenses and reaches mixed results.

This paper contributes to the existing literature by analyzing the overall effect of Internet diffusion on the most common categories of crime. To this end, I exploit the variation in broadband penetration and criminal activity in the US, in a panel data setup. The empirical challenge arises from the fact that Internet penetration is unlikely to be random with respect to the determinants of crime. As a consequence, standard Ordinary Least Squares ("OLS") regression would likely suffer from the omitted variable bias, and estimates would reflect a simple measure of association.

To more plausibly attach a causal interpretation to my results, I employ an Instrumental Variable ("IV") approach. My IV definition originates from the consideration that access to the Web requires being wired to a network. When the broadband Internet diffusion began, around the turn of the millennium, the telephone and cable TV companies were the first ones to offer access to their clients, because homes were already connected to copper wires (telephone) and coaxial cables (cable TV). Accordingly, connections using these pre-existing networks were preponderant in the early 2000s. Eventually, further technologies (with their respective networks) emerged and started penetrating the market, to the expenses of the older ones. My IV approach is based on the idea that the pre-Internet penetration pattern of the telephone and cable TV is correlated with broadband Internet diffusion and, most importantly, such correlation varies over time. On the basis of these considerations, I instrument broadband Internet penetration with a measure of early-times telephone and cable TV diffusion, multiplied by year indicators.

My IV estimates show that Internet diffusion is associated with a substantial and significant reduction in total and property crime rates. All else being equal, a broadband penetration increase by one additional line per 100 people is associated with a 1.4% decrease in the total crime rate and a 1.5% reduction in the property crime rate. Looking at specific offenses, I find that broadband Internet diffusion significantly affects each kind of property crime, while no significant impact is detectable on any category of violent offenses.

This design allows me to estimate the overall impact of Internet diffusion on crime – whereas I cannot disentangle between the content and time substitution channels. To better understand my findings, then, I rely on a theoretical framework derived from the results of previous research. Within this framework, the negative net impact of Internet on crime can be

best interpreted as reflecting a crime-reducing effect through the time substitution mechanism: time spent online crowds out alternative activities that are more likely to lead to crime. I cannot rule out that a positive, content-driven arousal effect is at play alongside with the negative time substitution impact. However, as far as total and property crimes are concerned, the latter more than outweighs the former so that, on net, the impact of the Internet is negative. On the other hand, the fact that no precise relationship is detected as regards (each category of) violent crime is consistent with the notion that the content-driven arousal effect is stronger on this type of offenses, and thus capable of neutralizing the negative time substitution impact. These results are in line with outcomes from previous non-Internet-specific studies.

I perform some robustness tests to challenge the validity of my IV results and the plausibility of their causal interpretation. To begin with, I run two falsification tests to check that outcomes are not driven by underlying state-specific trends in crime. Then, I verify the robustness of the estimated relationship by using a different source for Internet diffusion data. Finally, I replicate the analysis at a more disaggregated level. All in all, I find little cause for worry.

My research speaks to two policy-relevant debates. Since the beginning of the Internet era, policies have been put in place to subsidize the Web's infrastructure and universal deployment, by reason of the positive economic effect predicted for connected communities (for instance, Lehr *et al.*, 2006, ITU, 2012). However, policy-makers have been concerned about the possible negative effect of the Internet on some social aspects, including criminal activity. My results reveal that the overall effect of Internet diffusion on this outcome is negative and significant. This implies that, as far as crime is concerned, having the Internet available is better than not. Still, I find some evidence consistent with a positive content-driven effect, therefore it could be the case that appropriate measures to monitor and possibly restrict access to some types of content or users may further enhance the crime-reducing impact of the Web.

In addition, this research relates to the stream of literature investigating the crime drop experienced by many countries in the developed world since the early 1990s. The drop rebutted leading experts' predictions (for instance, Fox, 1996, and DiIulio, 1996), and has been object of a great deal of research aimed at pinning down its determinants. There is some consensus on a few factors that have concurred to falling crime rates (see Levitt, 2004). However, given the

complexity of the phenomenon, research on this topic is still ongoing. The present study relates to this literature and suggests that the Web has played a role in sustaining the crime drop.

The remainder of this Chapter is organized as follows. Section 2 provides a framework to help structure the discussion of the relationship between the Internet and crime, and reviews the relevant literature. Section 3 introduces the data and describes the empirical strategy. I present and discuss the results in Section 4. Section 5 reports some robustness checks and Section 6 concludes.

I.2 Framework and related literature

The Internet may affect crime outcomes via two channels, namely content and time substitution.

Being exposed to sensitive material (such as violence or pornography) could impact criminal behavior, in particular as far as violence is concerned (see Zimring and Hawkins, 1997). This impact could go in either direction. On the one hand, an arousal effect could make the audience more aggressive or prone to re-enact what they have seen online, thus generating a crime-enhancing result. On the other hand a catharsis effect may be deployed, so that potential criminals inhibit or satisfy their aggressive impulses through the consumption of sensitive content, and actual crimes are prevented. This dichotomy is clearly reflected, for instance, in the debate about pornography. Some scholars believe that sexually-explicit content raises the risk of sex crimes, because it triggers sexual arousal and aggression, degrades women and children to objects, and distorts social and individual norms (see for instance Zillman, 1971, Dworkin, 1981, Mackinnon, 1995). On the other side, it has been argued that the consumption of pornography may act as a substitute for sex crimes, and thus have a cathartic and crime-reducing role (see Posner, 1994). In conclusion, the impact of the Internet on crime deployed through the content channel is *a priori* ambiguous.

The Web has rapidly shifted from being an occasional tool to representing one of the main ways people interact, entertain themselves and work. A recent research shows that overall 73% of Americans go online on a daily basis, and one in five reports being connected “almost constantly” (see Pew Research Center, 2015). The time substitution mechanism is linked to the fact that the time we spend online comes at the expenses of some alternative occupation.

Wallsten (2013) uses data from the American Time Use Survey from 2003 to 2011 to estimate the crowd-out effects of the leisure time spent online. His results suggest that each minute of online leisure is correlated, among other things, with 0.29 fewer minutes on all other types of leisure (such as watching TV and video, offline socializing, and attending parties), 0.27 fewer minutes working, and 0.12 fewer minutes sleeping.

What matters for the purposes of understanding how the Web affects crime is whether the time spent online is more or less crime conducive than the alternative occupation that it crowds out. It could be that the displaced activities trigger crime at a lower rate than the time spent online, so that the Internet-crime relationship is positive. This may be the case, for example, if the Web is used by fellow criminals to coordinate or lure potential victims in a way that would not be possible otherwise. However, the story could also go the other way around if, say, the time spent online leads to earlier bedtimes and lower alcohol consumption as compared to bar attendance. In this case, spending time online reduces crime. The negative time substitution effect is also known in the literature as the voluntary incapacitation effect, whereby criminal interaction is reduced as long as potential offenders and victims voluntarily engage in some less crime-conducive occupation. In conclusion, the effect of the Internet on crime deployed through the time substitution channel is *a priori* ambiguous.

The mechanisms described above have been examined in previous empirical studies. The Internet-specific research on the issue is quite limited. It focuses on a specific kind of offenses – sex crimes – and reaches mixed results. Bhuller *et al.* (2011) investigate the impact of Internet use on sex crimes in Norway, in an IV setup. They find that Internet use is associated with a significant increase in sex offenses and interpret their results as reflecting an arousal effect. Kendall (2007) reaches opposite conclusions on the association between Internet use and rape by studying US state level data. He explains his results as expressing a catharsis effect, whereby criminals are substituting actual rape with pornography.

Non-Internet-specific empirical research has been carried on in related fields, finding substantial support for the arousal and the negative time substitution hypotheses, and very little evidence of a catharsis or positive time substitution impact.

Psychological research finds that subjects exposed to violent material tend to act more aggressively (see Anderson *et al.*, 2010; Anderson, Gentile, and Buckley, 2007). Results are typically derived from experiments in the lab, where a pool of subjects is randomly exposed to

some violent content, while the control group watches less violent material. Outcomes usually display significantly heightened cardiovascular activity and hostility measures associated with being exposed to more intense violence. These results suggest that there exists an arousal effect, at least as far as violent crimes are concerned.

Using real-world data, empirical research in economics has analyzed the effect on violent crime of being exposed to sensitive content over platforms other than the Internet. These studies find evidence of both the arousal and the negative time substitution effects. Results reveal that the latter more than outweighs the former, so that the net effect on crime is negative. Dahl and Della Vigna (2009) exploit time series variation in the violence level of blockbuster movies from 1995 to 2004 to study the overall short run effect of violent content on violent crime. They estimate that violent movie attendance reduces violent crime in the short term by between 1% and 2%. The authors interpret this finding as a negative time substitution effect: potential criminals who choose to go to the movie theatre substitute away from other activities that are more likely to lead to crime. They recognize that an arousal effect is also at play, evidenced by the finding that there are smaller reductions in violence after more violent movies as compared to mildly violent movies. Nonetheless, results show that the negative time substitution impact prevails. On a similar note, Cunningham, Engelstätter and Ward (2016) consider the effect of the increased weekly volume of violent video game sales on weekly violent crime. They conclude that violent video games may reduce crime by means of voluntary incapacitation, while their results show no support for the positive, content-driven effect.

In summary, each of the two mechanisms through which the Web can affect crime has an *a priori* ambiguous impact. Internet-specific research is rather limited and non-conclusive. Findings from non-Internet-specific studies suggest that, as far as the content channel is concerned, only the positive effect is material, in particular as far as violent crimes are concerned. On the other hand, regarding the time substitution mechanism, only the negative effect seems to be relevant. This paper relates and adds to the existing literature by empirically investigating the overall impact of Internet diffusion on crime.

I.3 Data and empirical strategy

I.3.1 Data

To investigate the link between Internet diffusion and crime evolution, I exploit the variation in broadband penetration and criminal activity in the US, in a panel data setup. I employ data on crime (the dependent variable), Internet penetration (the explanatory variable of interest), household telephone adoption and employment in the Cable TV sector (the instruments), as well as on a number of demographic and socio-economic controls. Data are organized in a panel covering a thirteen-year period (2000-2012) for 51 territories (50 states and the federal district). The information used in this study comes from several sources.

Crime data

State level annual data on crime occurrences are taken from the Uniform Crime Reports ("UCR") issued by the Federal Bureau of Investigation ("FBI"). These data have been used in mainstream crime literature, for instance by Kelly (2000) and Donohue and Levitt (2001). Data refer to seven felony offenses entailing both violent and property crimes. Violent crimes consist of murder and non-negligent manslaughter, forcible rape, robbery and aggravated assault; property crimes include burglary, larceny-theft and motor vehicle theft (see Table I.A-1 for definitions). The FBI selects these crime categories because they are the most likely to be reported and to occur with sufficient frequency to provide an adequate basis for comparison (see FBI, UCR Frequently Asked Questions). It is useful to stress that this classification does not include cybercrime which, despite being an intriguing phenomenon, is still quite uncommon. For instance, in the US in 2013 about 6 million larceny-thefts were reported, against 0.26 million cybercrime complaints (see FBI – UCR, 2013; FBI – IC3, 2013).

Crime rates per 100,000 residents represent the main dependent variables in the analysis. I use the logarithmic transformation because I expect the variation in crime to be proportional to its initial level. In a heterogeneity test, I use as dependent variables arrest rates for different combinations of age and gender. Data on arrests by age and sex are from the FBI's UCR. Finally, in a robustness check, I replicate the analysis at a more disaggregated level. Local data on crime are also from the FBI's UCR.

Internet data

The explanatory variable of interest must reflect Internet penetration. To the best of my knowledge, there are no state level data for the US available on a yearly basis on the diffusion of Internet technologies other than broadband. Broadband or high-speed Internet includes connections to end-user locations that deliver services at speeds exceeding 200 kilobits per second in at least one direction. The data I use picture the diffusion of broadband Internet only. Although this is not ideal, it shall be kept in mind that, starting from their introduction short before year 2000, broadband platforms have largely replaced other connections (NTIA, 2011; NTIA, 2013). In addition, the impact of broadband technologies has arguably been more intense than that of previous platforms (NTIA, 2004; NTIA, 2011), and it was only with the advent of broadband connections that the Internet reached preponderant levels of penetration in US homes. Thus, it does not seem unsound to approximate the effect of Internet diffusion in the period 2000-2012 using broadband penetration. I perform a robustness check using a different Internet data source (which includes all technologies but is only available on sporadic years) to get an indication as to whether this approximation is reasonable (see [section 5.2]).

In my baseline analysis, I use data on broadband diffusion from the Federal Communications Commission's ("FCC") Statistical Reports on Broadband Deployment. These have already been employed in the literature. For instance, Bellou (2015) uses them to examine the impact of broadband Internet penetration on marriage rates. The FCC's Statistical Reports on Broadband Deployment provide information on subscribership to Internet access services gathered through Form 477, which qualifying providers are required to file twice a year. Information includes data on the number of broadband or high-speed by type of user (residential or business) and per state. I focus on the residential segment because the crime-affecting mechanisms can be fully deployed only if the user can freely choose when to be online and which content to look at. Data on residential broadband penetration have been recorded since December 2000.[1] Because of this limitation, my analysis covers the period starting from 2000, although broadband deployment actually started at the end of the 1990s. I use for each year the second semester's report, portraying the situation as of the end of the year. Data are not available for

[1] Until the second release of 2004, Residential and Small Business are recorded as a single user type. Afterwards, the Residential and Business segments are recorded separately.

Hawaii in the period 2000-2005 and for Wyoming in 2000. My explanatory variable of interest is built as the number of residential broadband lines per 100 people in a certain state and year.

In a robustness check at both state and local level, I employ as an alternative source of information for Internet diffusion the Internet and Computer Use Supplement data collected by the US Census as part of its Current Population Survey ("CPS"). This source has been used in previous studies. For instance, Kuhn and Mansour (2014) exploited it to investigate the impact on unemployment duration of using the Internet within the job search strategy. CPS Internet data were only collected sporadically during the period of interest, which is why this is not my preferred source. I focus on a variable indicating whether the household has an Internet connection, regardless of the connection technology – the survey question is: "Does anyone in this household connect to the Internet from home? Yes/No". Over the period of interest, this question was asked in years 2000, 2001, 2003, 2007, 2009 and 2010. The fact that this variable is not broadband-specific allows me to get an indication as to whether the effects of broadband penetration can soundly approximate those of Internet diffusion at large.

Further data

I use data on state level residential telephone adoption and employment in the Cable TV sector in early years to build my instrumental variables. Data on residential telephone adoption in 1955 are from the 1956 City and County Data Book collected by the US Bureau of Census. Information is unavailable for Hawaii and Alaska. In a robustness test, I replicate the analysis at a more disaggregated level. Local data on residential telephone adoption in 1986 are from the CPS. State- and local-level data on employment in the Cable TV sector in 1990 are available by the US Bureau of Labor Statistics. I use US Standard Industrial Classification code 4841, referring to "Cable and other pay TV services".

I include a number of controls in my regressions. Data about states' characteristics (area, population, and age and race structure) are from the US Bureau of the Census. Figures on unemployment rates come from the US Bureau of Labor Statistics. Data on per capita personal income are available by the US Bureau of Economic Analysis. I consider the logarithmic transformation of this variable throughout the analysis.

Descriptive statistics

Table I.A-2 displays the mean and standard deviation in years 2000 and 2012 for the state level time-varying variables described above, as well as their percentage variation in mean over the said period. The pace of broadband Internet diffusion reflected by the FCC data is remarkable over the period 2000-2012, at more than 4000%. On the other hand, the increase in Internet penetration (all technologies) portrayed by the CPS data is less pronounced, as dial-up connections were already fairly diffused in 2000, when they started being replaced by broadband platforms. All crime rates have substantially fallen since the beginning of the millennium. In fact, much of the developed world has experienced a sharp drop in crime rates since the early 1990s.

I.3.2 Empirical strategy

I estimate the effect of broadband Internet penetration on crime rates using a panel that covers a thirteen-year period (2000-2012) for 51 territories (50 states and the federal district) in the US. I employ an IV approach, which is summarized by the following equations:

(I.1) $$y_{st} = \beta_0 + \beta_1 bbrate_{st} + \beta_2 X_{st} + \rho_t + \theta_s + \varepsilon_{st}$$

(I.2) $$bbrate_{st} = \gamma_0 + \gamma_1(\rho_t * phonerate_s) + \gamma_2(\rho_t * cableTV_s) + \gamma_3 X_{st} + \rho_t + \theta_s + \varepsilon_{st}$$

Where the subscripts s and t denote state and year respectively; y is the crime rate per 100,000 residents; $bbrate$ is the number of residential high-speed Internet lines per 100 residents; $phonerate$ is the number of residential phone lines per 100 residents in 1955; $cableTV$ is the number of Cable TV employees per 100,000 residents in 1990; ρ and θ are year and state fixed effects respectively; and X is a vector of demographic and socio-economic controls, namely: population density, unemployment rate, personal income per capita, percentage of Black or African American and percentage of males aged 15-24. I estimate the model above by Two-Stage Least Squares ("2SLS"), whereby Equation (I.1) represents the second stage and Equation (I.2) represents the first stage. I use the robust matrix variance estimator with standard errors clustered at the state level throughout the analysis.

Broadband diffusion is unlikely to be random with respect to the determinants of crime, even conditional on the model covariates. For example, states where cities are growing faster may experience faster broadband penetration and higher crime growth rates. As a consequence,

standard OLS estimation of the relationship between Internet diffusion and crime would likely suffer from the omitted variable bias, and estimates would reflect a measure of association rather than a causal effect. To more plausibly attach a causal interpretation to the estimation results, it is necessary to somehow filter the overall variation in broadband penetration and only use the part of it that is exogenous with respect to the determinants of crime evolution. The IV approach aims at implementing this intuitive strategy.

I instrument broadband Internet penetration with a measure of early-times telephone and cable TV diffusion, multiplied by year indicators. Specifically, for each state and year, I instrument $bbrate_{st}$, the fraction of population with a residential broadband Internet subscription, with $(\rho_t * phonerate_s)$ and $(\rho_t * cableTV_s)$, respectively the fraction of population which had a residential telephone in 1955 and the fraction of population which was employed in the Cable TV segment in 1990, multiplied by year indicators (for a similar approach, see Stevenson, 2006 and Bellou, 2015).

Broadband connection can be provided over different platforms, which entail different technologies. My IV definition originates from the consideration that, whichever the platform, access to the Web requires being wired to a network.

When the diffusion of broadband Internet began, around the turn of the millennium, the telephone and cable TV companies were the first ones to offer access, because homes were already connected to their networks. The Digital Subscriber Line ("DSL") is a wireline transmission technology that transmits data over the traditional copper telephone lines. The Cable Modem technology provides broadband through the same coaxial cables that deliver pictures and sound to the cable TV. Until the mid-2000s, these platforms faced virtually no competition accounting for the almost the totality of broadband connections. This period also displays the biggest share of variation in high-speed Internet penetration, as households quickly embraced this technology. Eventually, further platforms (and respective networks) such as Wireless and Fiber Optic emerged and started penetrating the market, to the expenses of the DSL and Cable Modem technologies. These dynamics are shown in Figure I.A-1, which portrays the residential broadband lines by technology, as well as the percentage variation in broadband diffusion, in the period 2000-2012.

The DSL and Cable Modem platforms account for a big share of broadband connections, and they rely on the pre-existing telephone and cable TV networks. Then, it can be argued that

the diffusion of the telephone and cable TV technologies before the Internet era is positively correlated with the subsequent broadband penetration, in a way that varies over time. In fact, until the late 2000s the correlation is expectedly stronger – with DSL and Cable Modem technologies accounting for the vast majority of connections – and eventually declines – as further platforms gain ground.

I argue that the time-varying correlation with the pre-Internet telephone and cable TV diffusion is relevant to the broadband rate evolution, even more so given that the relationship is stronger at the beginning of the sample, which is when the biggest share of variation in high-speed Internet penetration is concentrated.

I also argue that the part of variation in recent broadband penetration which is related to early-times telephone and cable TV deployment is exogenous with respect to the time-varying determinants of crime. In fact, this relationship is motivated by technological and business considerations (mainly, the existence of an adaptable previous network, whose costs were already sunk when the Internet era began) that are unlikely to bear any relevance to the time-varying factors affecting crime evolution. I do not dismiss that the (pre-Internet) adoption of telephone/cable TV (and the related recent broadband penetration pattern) may reflect some state-specific features – such as housing or cultural characteristics – which are stable over time and may well be correlated with (recent) crime patterns. However, the state fixed effects absorb this endogeneity.

In further support of the exogeneity requirement, the following considerations hold. Both the telephone and the cable TV technologies had virtually reached their full deployment by the turn of the millennium, when the broadband Internet diffusion began. This helps rule out possible direct effects of the instruments on recent crime patterns. Moreover, in choosing the reference years for the instruments, I try to comply with two constraints. On the one hand, I consider a moment in time that is “late enough” so that these technologies are already reasonably deployed. On the other hand, I make sure this moment is “early enough” so that network patterns do not incorporate any Internet-related consideration, and possible reverse effects of the dependent variable (crime rates in the period 2000-2012) on the instruments are prevented. The exogeneity condition cannot be tested directly. However, I run some robustness tests to enhance confidence that this requirement is met.

At a first visual inspection, data suggest that the telephone and Cable TV early deployment are indeed correlated with broadband diffusion, and that the strength of this link declines over time, as expected on the basis of the technological rationale underlying the relationship. Figure I.A-2 displays, for each year in the period 2000-2012, the actual state level broadband penetration rates plotted against those predicted by the share of residential telephone ownership in 1955 and the fraction of population employed in the Cable TV sector in 1990. In each year, the state dots are reasonably close to the 45 degree line (at different penetration levels), indicating the existence of a strong relationship. The predictive power of the telephone and Cable TV variables gets weaker over time. I further assess whether the instruments are relevant through the first stage estimation.

I.4 Results

I.4.1 The effect of broadband penetration on crime

I.4.1.1 First stage estimates

First stage estimation results are reported in Table I.A-3.[2] All instruments are positively correlated with broadband diffusion, implying that a higher level of penetration of the residential telephone and Cable TV technologies in the pre-Internet era is associated with greater broadband Internet diffusion.

The coefficients are highly significant starting from year 2003. Plausibly, in the initial period broadband was not diffused enough for a significant effect to be detected. As households rapidly adopted this technology and signed up for DSL or Cable Modem connections (see Figure I.A-1), the effect is precisely captured. Finally, around the end of the sample period, significance fades again, especially as far as the telephone adoption instrument is concerned. This is consistent with the fact that new broadband technologies emerged and gained ground against the older ones. All in all, the pattern of significance displayed by the estimated coefficients provides support to the technological motivation at the basis of my IV approach.

[2] For the sake of completeness, I present the first stage results using each of the two sets of instruments separately in Table I.B-1.

The F-statistic for a test on whether the excluded instruments are significant is close to 10, suggesting that the instruments are not weak.

I.4.1.2 Reduced form relation

Figure I.A-3 reports summary results from the reduced form estimation for the telephone adoption rate instruments. I obtained these coefficients by regressing crime rates directly against all instruments (telephone adoption*year and cable TV employment*year interactions) and further covariates. Extensive results are provided in Table I.B-2. The reduced form estimates display a negative and significant relation between telephone adoption rate and total and property crime rates. Importantly, the reduced form relation mirrors the pattern displayed by the first stage regression, whereby the link between the variables is not precisely estimated in the first years of the sample, it is significant in central years, and finally fades away at the end of the sample period. This parallel outline suggests that the reduced form association between crime and the instruments is indeed deployed through the technological rationale linking early diffusion of the telephone network and recent broadband Internet adoption.

The reduced form estimation detects no significant relationship between crime rates and the cable TV employment instruments, as displayed in Figure I.A-4.

I.4.1.3 Second stage estimates

Panel (A) in Table I.A-4 displays second stage estimates.[3] Results show a negative and significant effect of broadband Internet on total and property crime rates. Conversely, the estimated impact on violent crime is not significant. All else being equal, a broadband penetration increase by one additional line per 100 people is associated with a 1.4% decrease in total crime rates and a 1.5% reduction in property crime rates.

These coefficients should be interpreted as a local average treatment effect ("LATE"): by construction of my IV model, they represent the effect of broadband Internet diffusion on crime rates for that sub-population ("compliers") for which broadband penetrated according to the patterns of pre-existing telephone and Cable TV networks – that is, mainly, subscribers of DSL

[3] For the sake of completeness, I present the second stage results using each of the two sets of instruments separately in Table I.B-3.

and Cable Modem connections. I could not find socio-demographic data for this specific sub-population, therefore it is not possible to empirically get an indication of whether and how this group differs with respect to the average broadband subscriber. Still, there is no obvious reason to expect that a DSL subscriber should have substantially different characteristics with respect to, say, a Fiber subscriber.

As mentioned, exogeneity cannot be directly tested. At best, having more instruments than endogenous variables, the Hansen J test for overidentifying restrictions can be performed. This test assumes that at least one of the instruments is exogenous and verifies whether all instruments are jointly exogenous. Results from the Hansen J test are displayed in the last two rows of Table I.A-4. The joint null hypothesis is that the overidentifying restrictions are uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. The test does not reject the null hypothesis.

Panel (B) in Table I.A-4 displays outcomes from the OLS estimation of Equation (I.1). Again, results show a negative association between broadband diffusion and total and property crime rates. The estimated coefficients indicate that, all else being equal, a broadband penetration increase by one additional line per 100 people is associated with a 0.6% reduction in crime rates.

OLS estimated coefficients are about 2.3 times smaller in absolute value than 2SLS estimated coefficients. This suggests that the OLS approach does suffer from the omitted variable bias, as broadband diffusion is non-random with respect to the determinants of crime (conditional on other covariates). In particular, it seems to be the case that OLS estimates do not account for some heterogeneity that sets back the real effect of Internet penetration on crime rates. A consistent explanation for this bias would be, for instance, that states with faster growing cities record both a faster Internet penetration rate and slower crime declines.

I.4.2 Heterogeneity analyses

I.4.2.1 The effect of broadband diffusion on specific crimes

I study the effect of broadband diffusion on specific offenses. To do this, I perform my baseline IV approach using, as dependent variables, the natural logarithm of each offense’s rate per 100,000 people. Outcomes are reported in Table I.A-5.

Interestingly, results indicate that the impact of high-speed Internet is heterogeneous across property *versus* violent offenses. The estimated effects on each type of property felonies are significant. A broadband penetration increase by one additional line per 100 people is associated with reductions in burglary (-2.3%), larceny theft (-1%) and motor vehicle theft (-2.3%) rates. On the other hand, broadband penetration does not seem to have a significant effect on any of the violent offenses.

I compare my results on rape rates with those from previous studies on the link between Internet and sex crimes. Bhuller *et al.* (2013) use data referring to Norway in an IV setup and find that Internet use is associated with a significant increase in sex crimes, likely as a result of an arousal effect. On the other hand, Kendall (2007) reaches opposite conclusions on the association between Internet use and rape studying data on the US. My analysis suggests yet another result – that Internet diffusion does not significantly impact the occurrence of forcible rapes.

To help understand this discrepancy in results, it is useful to make some considerations. As regards the study by Bhuller *et al.*, the different geographic context may matter. Before the advent of the Internet, a legal ban on explicit pornography was in place in Norway – while it was *de facto* legalized and readily available in most of the US. This difference may at least in part explain the different consequences as regards sex crime outcomes. Concerning Kendall's (2007) study, the discrepancy in results likely arises from the different empirical strategy employed. The author uses an OLS estimation with fixed effects, thus assuming that the systematic determinants of Internet diffusion are time-invariant state characteristics. I relax this assumption by adopting an IV strategy.

I.4.2.2 The effect of broadband diffusion on different demographic groups

It has been suggested that the effect of Internet penetration may be heterogeneous across demographic groups. In particular, in a criminological study, Griffiths and Sutton (2013) propose that the time substitution effect may be stronger on the youth, who typically contribute more than proportionally to delinquency and are particularly keen on new technologies. In the words of the authors, "all the time spent online must equate to less time on the street leading to less potential offending time and a smaller population of available victims of violence and robbery" (Griffiths and Sutton, 2013, p. 19). Indeed, data indicate that young adults are in the vanguard of the

constantly connected, as 36% of 18-29-year-olds go online almost constantly (the corresponding figure is 21% for the US population at large) – see Pew Research Center, 2015.

I empirically test for heterogeneous effects by demographic group. An important data limitation is that there is no direct measure of the number of crimes committed by age group. Only when a crime is cleared by an arrest is it possible to attribute an age to the criminal. I employ FBI's UCR data on arrests by age and sex and replicate my IV analysis taking, as dependent variables, arrest rates for different combinations of age and gender. In particular, I consider the natural logarithm of the number of people arrested at age 0-14, 0-19, 0-24, 15-19, 15-24, 20+, 30+ as a percentage of people in that age range. I also consider the percentage of males arrested at age 0-19, 15-19 and 15-24 – this is because males are arrested at a considerably higher rate than females.

Second stage results are shown in Table I.A-6. Data do not seem to endorse the heterogeneous effect proposed by Griffiths and Sutton (2013). In fact, none of the estimated coefficients is significantly different from zero.

I.4.3 Discussion and economic significance

Results show a negative and significant relationship between broadband penetration and total and property crime rates. I estimate that a one percentage point increase in broadband penetration is associated with a 1.4% decrease in total crime rates and a 1.5% reduction in property crime rates. Since there were 3,226 offenses per 100,000 people – 2,854 of which were property offenses – reported in the US in 2012, a back-of-the-envelope calculation implies that one additional broadband line per 100 people translates into 45 fewer crimes per 100,000 people (43 of which are property crimes).

There are two channels through which such impact may deploy, namely content and time substitution. My design does not allow me to disentangle between these channels. To better understand my findings, then, I rely on previous literature which suggests that, as far as the content channel is concerned, only the positive effect is material. On the other hand, regarding the time substitution mechanism, it emerges that only the negative effect (or voluntary incapacitation) is relevant. Then, the crime-reducing impact of the Internet can be best interpreted as reflecting time substitution: time spent online crowds out alternative activities that are more likely to lead to crime. I cannot rule out that a positive arousal effect is at play alongside with the

negative time substitution one. However, as far as total and property offenses are concerned, the latter more than outweighs the former so that, on net, the impact of the Internet on crime is negative.

Previous literature suggests that the (positive) content mechanism is particularly material to violent crimes, while the (negative) time substitution one is relevant to both property and violent offenses. The fact that no precise relationship is detected with respect to (each category of) violent crime is consistent with the notion that the content-driven arousal effect on this type of offenses is more relevant, and capable of neutralizing the voluntary incapacitation effect.

Overall, my results are in line with the conclusions reached by the non-Internet-specific studies that assessed the effects on crime of technological platforms other than the Internet (Dahl and Della Vigna, 2009; Cunningham, Engelstätter and Ward, 2016).

I.5 Robustness

I.5.1 Falsification tests

I challenge the validity of my approach with two falsification tests. To begin with, I investigate whether different levels of diffusion in the pre-Internet telephone and cable TV networks across states are correlated with different crime dynamics in the period before the rise of the broadband technology. This outcome would suggest that my instruments are not exogenous and I am capturing a spurious correlation.

Broadband penetration occurred starting from the late 1990s – even though the analysis takes 2000 as the baseline year due to the lack of available data on the previous period. Specifically, I consider 1996 to be the year in which high-speed Internet penetration started because the literature (for instance, Bellou, 2015) indicates the 1996 Telecommunications Act – deregulating the broadcasting market and promoting competition in the telecommunications industry – as the turning point in broadband history.

This falsification test studies the association between the instruments and crime rates over the period 1990-2012. Summary results for the telephone adoption instrument are presented in Figure I.A-5. Extensive results are reported in Table I.B-4. Remarkably, the estimated

coefficients for total and property crime associated with telephone adoption start being consistently significant in 1996, and highly so (at the 1% level) in 1997.

As an additional test, following Bhuller *et al.* (2013), I include as further covariates in Equation (I.1) the interactions between baseline year (2000) exogenous covariates and either a linear time trend or time dummies. This falsification test allows crime evolution to be related to state-specific trends in the exogenous covariates. The purpose of this analysis is to check whether the significant effect detected in my baseline analysis holds or is absorbed by such trends. Equation (I.1) becomes:

(I.3) $y_{st} = \beta_0 + \beta_1 bbrate_{st} + \beta_2 X_{st} + t \sum_j \lambda_j x_{(s,2000),j} + \rho_t + \theta_s + \varepsilon_{st}$, with a time trend;

(I.4) $y_{st} = \beta_0 + \beta_1 bbrate_{st} + \beta_2 X_{st} + \rho_t \sum_j \lambda_j x_{(s,2000),j} + \theta_s + \varepsilon_{st}$, with time dummies.

Where x_j, $j = (1, \dots, J)$, are the covariates I am using in my baseline model.

Table I.A-7 displays outcomes from the 2SLS estimation of Equation (I.3) – panel (A) – and Equation (I.4) – panel (B). Results show that the negative and significant relationship between broadband diffusion and total and property crime rates is not accounted for by underlying state-specific trends in the determinants of crime. In fact, the estimated coefficients for total and property crime are still negative and significant. Point estimates from the estimation of Equation (I.3) are very close to my baseline results, while those from Equation (I.4) are a bit larger in absolute value.

I.5.2 Analysis using CPS Internet data

I further test my results by replicating the analysis with a different source of data on Internet diffusion. I use the Internet and Computer Use Supplement data collected by the US Census as part of the CPS. This source has been widely used in previous literature on the effects of Internet penetration on socio-economic outcomes (for instance, Kendall, 2007; Kuhn and Mansour, 2014). I focus on a variable indicating whether the household has an Internet connection, regardless of the connection technology – the survey question is: “Does anyone in this household connect to the Internet from home? Yes/No”. This variable is available in six years over the period 2000-2012 (namely: 2000, 2001, 2003, 2007, 2009 and 2010), and has been used in previous related literature (for instance, Kendall, 2007).

Equation (I.1) and Equation (I.2) become, respectively:

(I.5) $$y_{st} = \beta_0 + \beta_1 HHint\_rate_{st} + \beta_2 X_{st} + \rho_t + \theta_s + \varepsilon_{st}$$

(I.6) $$HHint\_rate_{st} = \gamma_0 + \gamma_1(\rho_t * phonerate_s) + \gamma_2(\rho_t * cableTV_s) + \gamma_3 X_{st} + \rho_t + \theta_s + \varepsilon_{st}$$

Where $HHint\_rate_{st}$ is the natural logarithm of the percentage of households that have an Internet connection in state *s* and year *t*.

The purpose of this specification check is twofold. Firstly, since the CPS Supplement has been widely used in previous literature, it is sensible to replicate my analysis using this source to check whether my baseline results are confirmed. Secondly, this test provides some indication as to whether the effect of the diffusion of broadband technologies can be soundly used to approximate that of the Internet diffusion at large.

Such approximation is necessary because of data limitation issues, as overall Internet diffusion data are only available on sporadic years. Still, in my baseline analysis, I have assumed that it is sound, based on the facts that (i) starting from their introduction around year 2000, broadband platforms have largely replaced dial-up connections (NTIA, 2011; NTIA, 2013); (ii) the impact of broadband technologies has arguably been more intense than that of dial-up connections (NTIA, 2004; NTIA, 2011), and (iii) it was only with the introduction of broadband technologies that the Internet reached preponderant levels of penetration in US homes.

In this specification check, I use the CPS data for the available years and consider a variable that covers all types of Internet connections. If the analysis were to generate substantially different results with respect to my baseline outcomes, this could imply that the effect on crime of broadband technologies penetration is considerably different with respect to that of the Internet at large.

Panel (C) in Table I.A-7 presents the second stage results. For completeness, I show in panel (D) the outcomes from OLS estimation of Equation (I.5). Even using a different source of data on Internet diffusion, this is estimated to have a negative and significant impact on total (-1.6%) and property (-1.7%) crime rates. The IV-estimated coefficients are close in magnitude to my baseline results. These findings provide some evidence that the effect on crime of broadband technologies penetration in the period 2000-2012 can be soundly used to approximate that of the Internet at large.

I.5.3 Local level analysis

States are huge aggregates and a great deal of variation goes on within their boundaries. Because of this, a concern with my baseline analysis is that the estimation method may not be appropriate, as too big entities are being considered. To address this issue, I replicate the analysis at a more disaggregated level.

To the best of my knowledge, the only source of Internet data at the local level is represented by the CPS Internet and Computer Use Supplements, which report data at the Metropolitan Statistical Area ("MSA") level. MSAs are geographic entities institutionally delineated for statistical purposes and consisting of one or more counties. Caution shall be paid because CPS data are not representative for all parts of the US at the MSA level.

I focus on a variable indicating whether the household has an Internet connection, regardless of the technology – the survey question is: "Does anyone in this household connect to the Internet from home? Yes/No". This variable is available in six years over the period 2000-2012 (namely: 2000, 2001, 2003, 2007, 2009 and 2010), and has been used in previous related literature (Kendall, 2007).

Local data on crime are available by the FBI's UCR at the county level. They include a coverage indicator (increasing from 0 to 100), which represents a measure of data quality for each county/year combination. I set a minimum threshold for acceptable data quality at 90/100 (using different floor thresholds does not significantly impact results). I transform county-level data into MSA-level data and combine them with the Internet diffusion data in a yearly (six years with gaps), MSA-level panel. MSAs are re-delineated periodically. I focus on the 112 MSAs for which data on the dependent and control variables are available in all sample years. These represent about one third of the total number of MSAs.

Table I.A-8 displays the mean and standard deviation for the MSA-level time-varying variables in years 2000 and 2010, as well as their percentage variation in mean over such period. Data are in line with state level figures presented in Table I.A-2 (even though those statistics refer to years 2000 and 2012). Data confirm that crime rates have been falling while the Internet was expanding. Crime rates are a bit higher at the MSA level in comparison with the state level. This is likely due to the fact that MSAs are by definition densely populated areas.

For this local level study, I implement both an IV and an OLS setup that parallel the state level analysis. As instrumental variables, I use residential telephone adoption rate in 1986 (the

first year in which information is reported for a considerable number of MSAs) and the share of population employed in the Cable TV sector in 1990, at the MSA level. Equation (I.1) and Equation (I.2) become, respectively:

(I.7) $$y_{st} = \beta_0 + \beta_1 HHint\_rate_{st} + \rho_t + \theta_s + \varepsilon_{st}$$

(I.8) $$HHint\_rate_{st} = \gamma_0 + \gamma_1(\rho_t * phonerate_s) + \gamma_2(\rho_t * cableTV_s) + \rho_t + \theta_s + \varepsilon_{st}$$

Where the subscript *s* denotes MSA and θ_s are MSA fixed effects. I use the robust matrix variance estimator with standard errors clustered at the MSA level.

In contrast with respect to the state level outcomes, my instruments do not seem to work well in the local setup. In fact, first stage results in Table I.A-9 show that there is no significant association between the instruments and the endogenous variable in any of the sample years, implying a lack of relevance of the instruments at the local level.[4]

Panel (A) in Table I.A-10 presents the second stage results from the IV model,[5] whereas panel (B) displays the outcomes from OLS estimation of Equation (I.7). All estimated coefficients are negative, and point estimates are not far in magnitude from those obtained in my baseline analysis. However, possibly as a result of the lack of relevance of the instruments, only the OLS estimated coefficients for total and property crime are significant.

Although no considerations on causality can be drawn from the local analysis, outcomes from this specification lend support to the finding of a negative and significant association between Internet diffusion and crime, in line with my state level results.

I.6 Conclusions

The Internet may affect crime outcomes through two mechanisms – content and time substitution – and each of them has an *a priori* ambivalent impact. I empirically investigate the overall effect of Internet diffusion on the most common categories of crime in the US, employing an IV strategy in a panel data setting.

---

[4] For the sake of completeness, I present the first stage results using each of the two sets of instruments separately in Table I.B-5.

[5] For the sake of completeness, I present the second stage results using each of the two sets of instruments separately in Table I.B-6.

My findings show a negative and significant relationship between broadband penetration and total and property crime rates. On the other hand, no precise association is detected as regards violent crimes. I perform a number of robustness checks to challenge the validity of the estimated relationship, finding little cause for worry.

To better understand my findings, I rely on a theoretical framework derived from the results of previous research. Within this framework, the negative net impact of Internet on crime can be best interpreted as reflecting a crime-reducing effect through time substitution: time spent online crowds out alternative activities that are more likely to lead to crime. I cannot rule out that a positive, content-driven arousal effect is at play alongside the negative time substitution one. However, as far as total and property offenses are concerned, the latter more than outweighs the former so that, on net, the Internet has a crime-reducing impact. The fact that no precise relationship is detected with respect to (each category of) violent crime is consistent with the hypothesis that the content-driven arousal effect on this type of offenses is stronger, and thus capable of neutralizing the time substitution impact.

This study speaks to the ongoing debate as to the far-reaching social impacts of the Web. My results reveal that the overall effect of Internet diffusion on crime is negative and significant. This implies that, as far as this outcome is concerned, having the Internet available is better than not. Still, my findings do not dismiss that some crime-increasing arousal effect may be at play. Then, further research is needed to assess whether monitoring or restricting some types of content or users may further enhance the crime-reducing impact of the Web. Finally, this paper contributes to the stream of literature investigating the crime drop experienced by the US since the early 1990s. My findings suggest that Internet diffusion should be enlisted among the factors that have jointly concurred to sustaining the decline.

Appendix I.A Figures and tables

TABLE I.A-1: FBI'S OFFENSE DEFINITIONS

Offense	Type	Definition
Murder and non-negligent manslaughter	Violent	The willful (non-negligent) killing of one human being by another. Deaths caused by negligence, attempts to kill, assaults to kill, suicides, and accidental deaths are excluded. The program classifies justifiable homicides separately and limits the definition to: (1) the killing of a felon by a law enforcement officer in the line of duty; or (2) the killing of a felon, during the commission of a felony, by a private citizen.
Forcible rape	Violent	The carnal knowledge of a female forcibly and against her will. Rapes by force and attempts or assaults to rape, regardless of the age of the victim, are included. Statutory offenses (no force used—victim under age of consent) are excluded.
Robbery	Violent	The taking or attempting to take anything of value from the care, custody, or control of a person or persons by force or threat of force or violence and/or by putting the victim in fear.
Aggravated assault	Violent	An unlawful attack by one person upon another for the purpose of inflicting severe or aggravated bodily injury. This type of assault usually is accompanied by the use of a weapon or by means likely to produce death or great bodily harm. Simple assaults are excluded.
Burglary	Property	The unlawful entry of a structure to commit a felony or a theft. Attempted forcible entry is included.
Larceny-theft	Property	The unlawful taking, carrying, leading, or riding away of property from the possession or constructive possession of another (except motor vehicle theft). Examples are thefts of bicycles, motor vehicle parts and accessories, shoplifting, pocket-picking, or the stealing of any property or article that is not taken by force and violence or by fraud. Attempted larcenies are included. Embezzlement, confidence games, forgery, check fraud, etc., are excluded.
Motor vehicle theft	Property	The theft or attempted theft of a motor vehicle. A motor vehicle is self-propelled and runs on land surface and not on rails. Motorboats, construction equipment, airplanes, and farming equipment are specifically excluded from this category.

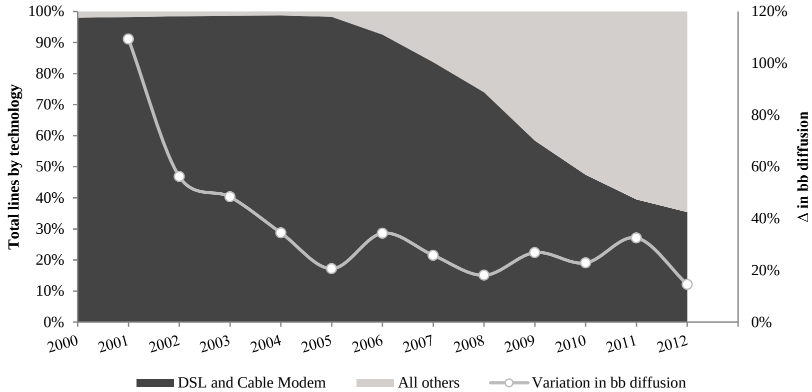
Source: FBI, UCR Frequently Asked Questions.

TABLE I.A-2: DESCRIPTIVE STATISTICS, 2000 AND 2012

Variable	2000		2012		Change (%)
	Mean	SD	Mean	SD	
Total crime rate	4,052.73	1,040.01	3,225.61	737.87	-20.41
Violent crime rate	443.33	241.14	371.27	178.66	-16.25
Property crime rate	3,609.40	868.12	2,854.33	606.76	-20.92
Residential broadband lines per 100 people – from FCC's reports	1.58	0.94	66.95	6.77	4,137.34
Households with Internet connection (%) – from the CPS	45.11	6.70	73.96(*)	6.11(*)	64
Black or Afro-Americans (%)	11.15	11.91	11.54	10.99	3.50
Male population aged 15-24 (%)	7.27	0.64	7.17	0.39	-1.38
Population density (/sqmi)	315.78	1,172.75	345.24	1,294.97	9.33
Personal income per capita ($)	29,441.35	4,804.47	44,014.27	7,956.04	49.50
Unemployment rate (%)	3.91	0.93	7.37	1.71	88.49

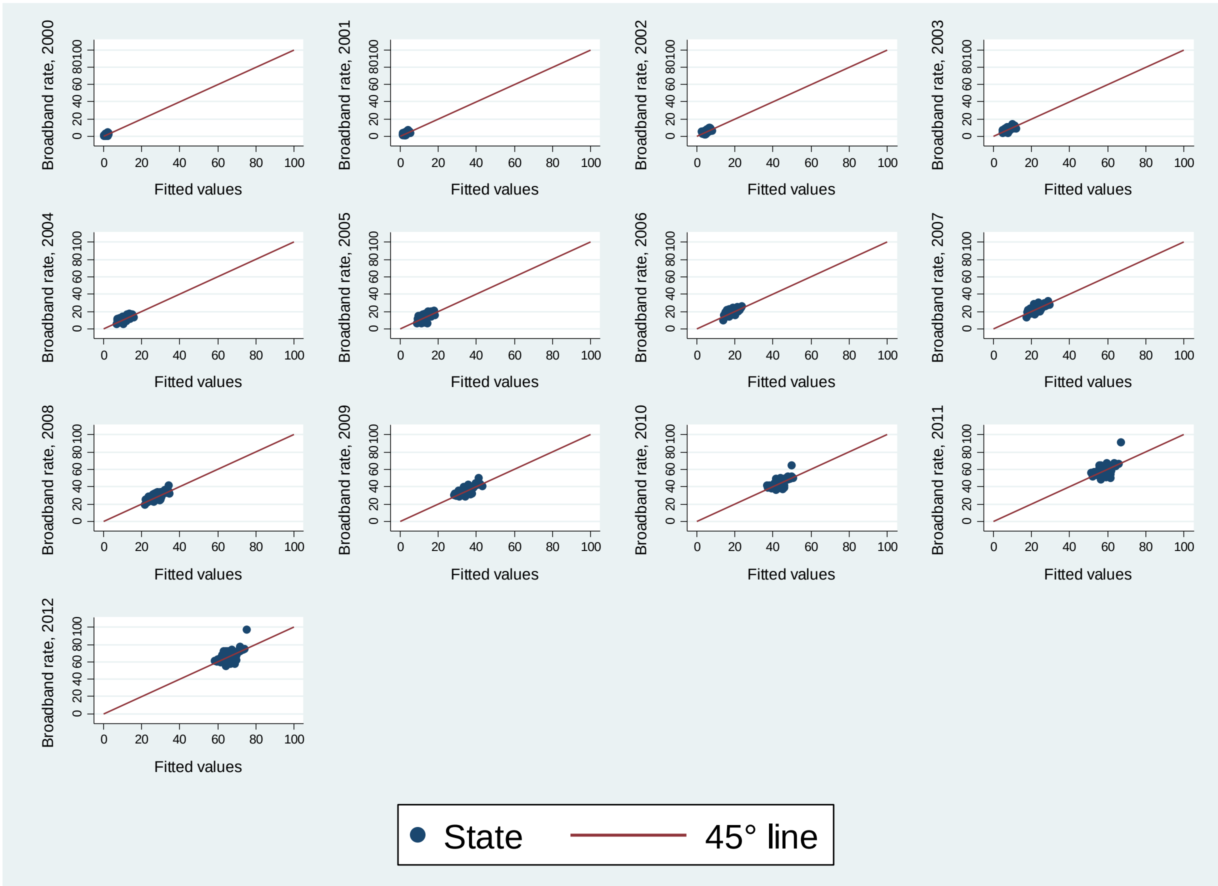
Notes: Summary statistics for years 2000 and 2012 are generated from yearly figures at the state level. Crime rates are per 100,000 people. (*) These values refer to year 2010.

FIGURE I.A-1: RESIDENTIAL BROADBAND INTERNET LINES BY TECHNOLOGY AND VARIATION IN BROADBAND DIFFUSION, 2000-2012



Notes: The main vertical axis shows the share of residential broadband Internet connections accounted for by the DSL and Cable Modem platforms vs all other technologies in years 2000-2012. The secondary vertical axis measures the year-on-year percentage change in the number of residential broadband lines per 100 people in the period 2001-2012.

FIGURE I.A-2: RESIDENTIAL BROADBAND LINES (PER 100 PEOPLE): ACTUAL *VS* PREDICTED BY RATE OF RESIDENTIAL TELEPHONE ADOPTION (1955) AND RATE OF EMPLOYMENT IN THE CABLE TV SECTOR (1990)



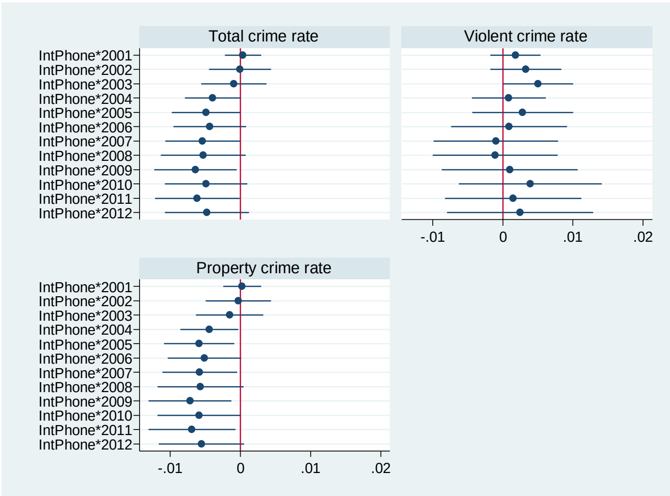
Notes: Each of the figures above refers to one year in the period 2000-2012. Separately for each year, I regress the state-level residential broadband Internet penetration (number of lines per 100 people) against the state-level share of residential telephone ownership in 1955 and the state-level fraction of population employed in the Cable TV sector in 1990. Based on these regressions, I obtain the linear prediction of the state-level residential broadband Internet penetration. In each figure, the horizontal axis measures the predicted residential broadband Internet penetration. The vertical axis measures the actual residential broadband Internet penetration. Each dot corresponds to one state.

TABLE I.A-3: FIRST STAGE ESTIMATION

DV: Broadband lines rate	*Phone adoption rate	*Rate of employment in Cable TV sector
(year==2001)*	0.0217	0.0241*
	[0.0209]	[0.0127]
(year==2002)*	0.044	0.0364***
	[0.0344]	[0.0138]
(year==2003)*	0.1551***	0.0525***
	[0.0578]	[0.0186]
(year==2004)*	0.2083***	0.0665***
	[0.0630]	[0.0247]
(year==2005)*	0.2848***	0.0600**
	[0.0811]	[0.0236]
(year==2006)*	0.2885***	0.0491*
	[0.0672]	[0.0277]
(year==2007)*	0.3023***	0.0886***
	[0.0715]	[0.0183]
(year==2008)*	0.3103***	0.0967***
	[0.0734]	[0.0142]
(year==2009)*	0.2699***	0.1240***
	[0.0667]	[0.0149]
(year==2010)*	0.2085**	0.1084***
	[0.0846]	[0.0235]
(year==2011)*	0.2323*	0.0851**
	[0.1204]	[0.0359]
(year==2012)*	0.1905*	0.1222***
	[0.0993]	[0.0362]
Year effects		Yes
Other controls		Yes
N		636
F-test for IV		9.7528
F-test for IV - Prob > F		0

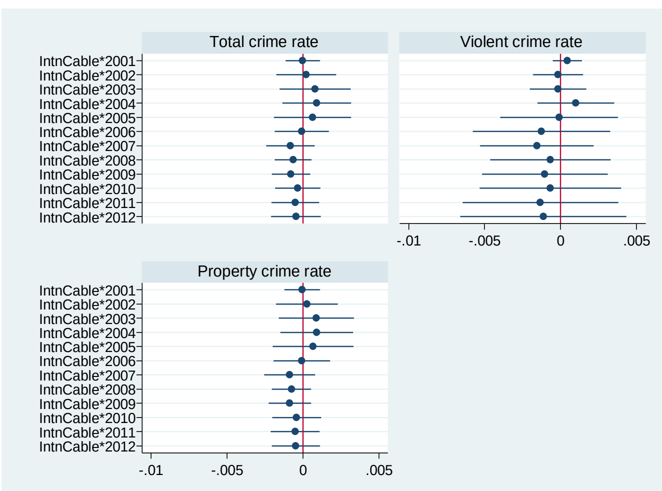
Notes: Observations are yearly and at the state level for the period 2000-2012. The coefficients for the two sets of instruments (year phone adoption rate and year* rate of employment in Cable TV sector) are estimated from the same model, but reported next to each other for convenience. The estimates come from ordinary least squares regressions with state and year fixed effects. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

FIGURE I.A-3: REDUCED FORM ESTIMATION, $\rho_t * phonerate_s$



Notes: I use yearly observations at the state level for the period 2000-2012. I regress crime rates against two sets of regressors (year phone adoption rate and year* rate of employment in Cable TV sector) jointly and display the estimated coefficients for the former set. Estimates come from ordinary least squares regressions with state and year fixed effects, with robust standard errors clustered at the state level. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. The coefficients are displayed with their respective 95% confidence intervals.*

FIGURE I.A-4: REDUCED FORM ESTIMATION, $\rho_t * cableTV_s$



Notes: I use yearly observations at the state level for the period 2000-2012. I regress crime rates against two sets of regressors (year phone adoption rate and year* rate of employment in Cable TV sector) jointly and display the estimated coefficients for the latter set. Estimates come from ordinary least squares regressions with state and year fixed effects, with robust standard errors clustered at the state level. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. The coefficients are displayed with their respective 95% confidence intervals.*

TABLE I.A-4: SECOND STAGE AND OLS ESTIMATION

DV: Crime rate	(A) 2SLS			(B) OLS		
	Total	**Violent**	**Property**	**Total**	**Violent**	**Property**
Broadband lines rate	-0.0141***	-0.0094	-0.0153***	-0.0060**	-0.0032	-0.0065**
	[0.0051]	[0.0114]	[0.0051]	[0.0026]	[0.0040]	[0.0027]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
N	636	636	636	656	656	656
Hansen J statistic	24.757	27.982	23.777			
p-value	0.3629	0.2165	0.4163			

Notes: Observations are yearly and at the state level for the period 2000-2012. The estimates in panel (A) come from 2SLS regressions with state and year fixed effects, whereby I use the two sets of IVs (year phone adoption rate and year* rate of employment in Cable TV sector) to instrument for broadband Internet penetration. The estimates in panel (B) come from ordinary least squares regressions with state and year fixed effects, whereby I regress the dependent variables directly against broadband Internet penetration. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE I.A-5: SECOND STAGE ESTIMATION – SPECIFIC CRIMES

DV: Crime rate	Violent crimes				Property crimes		
	Murder	Forcible rape	Robbery	Aggravated assault	Burglary	Larceny theft	Motor vehicle theft
Broadband lines rate	0.0026	-0.0121	-0.002	-0.0129	-0.0235***	-0.0100**	-0.0231**
	[0.0130]	[0.0106]	[0.0088]	[0.0149]	[0.0074]	[0.0050]	[0.0101]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	636	636	636	636	636	636	636

Notes: Observations are yearly and at the state level for the period 2000-2012. The estimates come from 2SLS regressions with state and year fixed effects, whereby I use the two sets of IVs (year phone adoption rate and year* rate of employment in Cable TV sector) to instrument for broadband Internet penetration. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

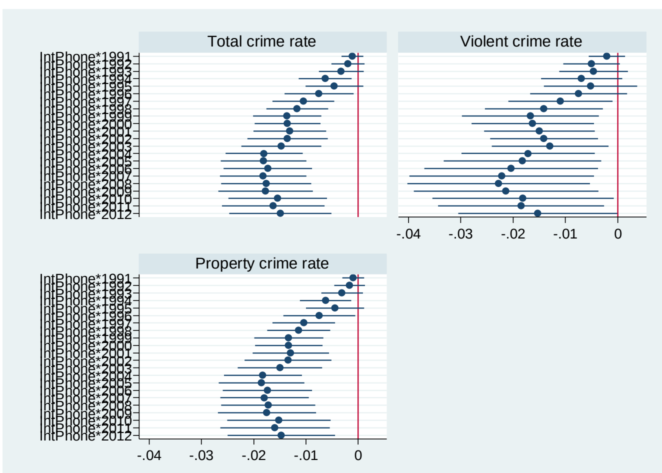
TABLE I.A-6: SECOND STAGE ESTIMATION – ARREST RATES BY DEMOGRAPHIC GROUP

DV: Arrest rate per demographic group	Age 0-14	Age 0-19	Age 0-24	Age 15-19	Age 15-24
Broadband lines rate	-0.012	0	-0.0014	0.0017	0.0004
	[0.0195]	[0.0159]	[0.0174]	[0.0165]	[0.0176]
Year effects	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes
N	621	621	621	621	621

DV: Arrest rate per demographic group	Age 20+	Age 25+	Age 30+	Age 0-19 Males	Age 15-19 Males	Age 15-24 Males
Broadband lines rate	-0.0105	-0.0143	-0.0155	-0.0009	0.0007	0.0003
	[0.0214]	[0.0225]	[0.0228]	[0.0162]	[0.0166]	[0.0179]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
N	621	621	621	621	621	621

Notes: Observations are yearly and at the state level for the period 2000-2012. The estimates come from 2SLS regressions with state and year fixed effects, whereby I use the two sets of IVs (year phone adoption rate and year* rate of employment in Cable TV sector) to instrument for broadband Internet penetration. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. The estimation was replicated without the latter control: results remain unchanged. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

FIGURE I.A-5: FALSIFICATION TEST, $\rho_t * phonerate_s$, 1990-2012



Notes: I use yearly observations at the state level for the period 1990-2012. I regress crime rates against two sets of regressors (year phone adoption rate and year* rate of employment in Cable TV sector) jointly and display the estimated coefficients for the former set. Estimates come from ordinary least squares regressions with state and year fixed effects, with robust standard errors clustered at the state level. Other controls include population density, unemployment rate, personal income per capita, percentage of males aged 15-24. The coefficients are displayed with their respective 95% confidence intervals.*

TABLE I.A-7: ROBUSTNESS TESTS WITH TIME INTERACTIONS AND RESULTS USING INTERNET DATA FROM THE CPS

DV: Crime rate	(A) Time trend			(B) Time dummies		
	Total	Violent	Property	Total	Violent	Property
Broadband lines rate	-0.0154***	-0.0217**	-0.0153***	-0.0204***	-0.0200**	-0.0205***
	[0.0051]	[0.0106]	[0.0051]	[0.0054]	[0.0094]	[0.0057]
Year effects	Yes	Yes	Yes	No	No	No
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
N	636	636	636	636	636	636
DV: Crime rate	**(C) CPS data - 2SLS**			**(D) CPS data - OLS**		
	Total	Violent	Property	Total	Violent	Property
HH has Internet (%)	-0.0157**	-0.0092	-0.0170**	-0.0041**	-0.0029	-0.0043**
	[0.0075]	[0.0089]	[0.0077]	[0.0019]	[0.0027]	[0.0020]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
N	294	294	294	306	306	306

Notes: Observations are yearly and at the state level for the period 2000-2012. In panels (A) the estimates come from 2SLS regressions with state and year fixed effects, whereby I use the two sets of IVs (year phone adoption rate and year* rate of employment in Cable TV sector) to instrument for broadband Internet penetration. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24 and the interactions between baseline year (2000) exogenous covariates and a linear time trend. In panel (B) the estimates come from 2SLS regressions with state fixed effects, whereby I use the two sets of IVs to instrument for broadband Internet penetration. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24 and the interactions between baseline year (2000) exogenous covariates and year dummies. In panel (C) the estimates come from 2SLS regressions with state and year fixed effects, whereby I use the two sets of IVs to instrument for the percentage of households with an Internet connection. In panel (D) the estimates come from ordinary least squares regressions with state and year fixed effects, whereby I regress the dependent variables directly against the percentage of households with an Internet connection. In panels (C) and (D) other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. "HH" stands for household. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE I.A-8: DESCRIPTIVE STATISTICS – MSA-LEVEL DATA, 2000 AND 2010

Variable	2000		2010		Change (%)
	Mean	SD	Mean	SD	
Total crime rate	4,357.87	2,023.63	3,745.89	1,979.23	-14.04
Violent crime rate	453.05	262.11	417.92	271.81	-7.75
Property crime rate	3,904.83	1,813.73	3,327.97	1,745.60	-14.77
HH with Internet connection (%) – from the CPS	47.94	10.76	74.95	9.58	56.34

Notes: Summary statistics for years 2000 and 2010 are generated from yearly figures at the MSA level. Crime rates are per 100,000 people. "HH" stands for household.

TABLE I.A-9: FIRST STAGE ESTIMATION – MSA-LEVEL ANALYSIS

DV: HHs with Internet connection (%)	*Phone adoption rate	*Rate of employment in Cable TV sector
(year==2001)*	1.871	-0.0359
	[5.4539]	[0.0371]
(year==2003)*	6.8351	-0.0406
	[5.6994]	[0.0382]
(year==2007)*	-4.361	-0.0249
	[4.9199]	[0.0379]
(year==2009)*	4.6845	-0.0074
	[4.9342]	[0.0365]
(year==2010)*	-2.5501	-0.0343
	[6.2585]	[0.0398]
Year effects		Yes
Other controls		No
N		540
F-test for IV		1.16
F-test for IV - Prob > F		0.2382

Notes: Observations are yearly and at the MSA level for years 2000, 2001, 2003, 2007, 2009 and 2010. The coefficients for the two sets of instruments (year phone adoption rate and year* rate of employment in Cable TV sector) are estimated from the same model, but reported next to each other for convenience. The estimates come from ordinary least squares regressions with MSA and year fixed effects. No other controls are included. "HH" stands for household. Robust standard errors (in brackets) are clustered at the MSA level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE I.A-10: SECOND STAGE AND OLS ESTIMATION – MSA-LEVEL ANALYSIS

DV: Crime rate	(A) 2SLS			(B) OLS		
	Total	Violent	Property	Total	Violent	Property
HH has Internet (%)	-0.0118	-0.0057	-0.0126	-0.0046**	-0.0034*	-0.0047**
	[0.0120]	[0.0137]	[0.0120]	[0.0019]	[0.0020]	[0.0020]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	No	No	No	No	No	No
N	531	531	531	658	658	658

Notes: Observations are yearly and at the MSA level for years 2000, 2001, 2003, 2007, 2009 and 2010. The estimates in panel (A) come from 2SLS regressions with MSA and year fixed effects, whereby I use the two sets of IVs (year phone adoption rate and year* rate of employment in Cable TV sector) to instrument for the percentage of households with an Internet connection. The estimates in panel (B) come from ordinary least squares regressions with MSA and year fixed effects, whereby I regress the dependent variables directly against the percentage of households with an Internet connection. No other controls are included. "HH" stands for household. Robust standard errors (in brackets) are clustered at the MSA level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

Appendix I.B Additional figures and tables

TABLE I.B-1: FIRST STAGE ESTIMATION USING THE TWO SETS OF INSTRUMENTS SEPARATELY

DV: Residential broadband lines rate		
Interaction	IV=Phone adoption rate	IV=Rate of employment in Cable TV sector
(year==2001)*	0.0339	0.0298**
	[0.0225]	[0.0146]
(year==2002)*	0.0608*	0.0454***
	[0.0363]	[0.0158]
(year==2003)*	0.1794***	0.0693***
	[0.0564]	[0.0204]
(year==2004)*	0.2427***	0.0864***
	[0.0615]	[0.0263]
(year==2005)*	0.3064***	0.0842***
	[0.0811]	[0.0243]
(year==2006)*	0.3032***	0.0889***
	[0.0685]	[0.0308]
(year==2007)*	0.3407***	0.1252***
	[0.0743]	[0.0234]
(year==2008)*	0.3526***	0.1232***
	[0.0765]	[0.0163]
(year==2009)*	0.3239***	0.1352***
	[0.0772]	[0.0181]
(year==2010)*	0.2523***	0.1141***
	[0.0901]	[0.0242]
(year==2011)*	0.2678**	0.0948***
	[0.1197]	[0.0309]
(year==2012)*	0.2506**	0.1296***
	[0.1019]	[0.0311]
Year effects	Yes	Yes
Other controls	Yes	Yes
N	636	656
F-test for IV	5.16	10.91
F-test for IV - Prob > F	0	0

*Notes: Observations are yearly and at the state level for the period 2000-2012. The estimates come from ordinary least squares regressions with state and year fixed effects. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE I.B-2: REDUCED FORM ESTIMATION

DV: Crime rate	(A) Total		(B) Violent		(C) Property	
Interaction	*Phone adoption rate	*Rate of employment in Cable TV sector	*Phone adoption rate	*Rate of employment in Cable TV sector	*Phone adoption rate	*Rate of employment in Cable TV sector
(year==2001)*	0.0004	0	0.0018	0.0004	0.0002	0
	[0.0013]	[0.0006]	[0.0018]	[0.0005]	[0.0013]	[0.0006]
(year==2002)*	-0.0001	0.0002	0.0033	-0.0002	-0.0003	0.0003
	[0.0022]	[0.0010]	[0.0025]	[0.0008]	[0.0023]	[0.0010]
(year==2003)*	-0.0009	0.0008	0.0050**	-0.0002	-0.0015	0.0009
	[0.0023]	[0.0012]	[0.0025]	[0.0009]	[0.0024]	[0.0012]
(year==2004)*	-0.0039*	0.0009	0.0009	0.001	-0.0044**	0.0009
	[0.0020]	[0.0011]	[0.0026]	[0.0013]	[0.0020]	[0.0012]
(year==2005)*	-0.0049*	0.0006	0.0028	-0.0001	-0.0059**	0.0007
	[0.0024]	[0.0013]	[0.0036]	[0.0019]	[0.0025]	[0.0013]
(year==2006)*	-0.0044*	-0.0001	0.0009	-0.0013	-0.0051*	-0.0001
	[0.0026]	[0.0009]	[0.0041]	[0.0023]	[0.0026]	[0.0009]
(year==2007)*	-0.0054**	-0.0008	-0.001	-0.0016	-0.0058**	-0.0009
	[0.0027]	[0.0008]	[0.0044]	[0.0019]	[0.0027]	[0.0008]
(year==2008)*	-0.0053*	-0.0006	-0.0011	-0.0007	-0.0057*	-0.0008
	[0.0030]	[0.0006]	[0.0044]	[0.0020]	[0.0031]	[0.0006]
(year==2009)*	-0.0064**	-0.0008	0.001	-0.001	-0.0072**	-0.0009
	[0.0029]	[0.0006]	[0.0048]	[0.0021]	[0.0029]	[0.0007]
(year==2010)*	-0.0049*	-0.0004	0.0039	-0.0007	-0.0059*	-0.0004
	[0.0029]	[0.0007]	[0.0051]	[0.0023]	[0.0030]	[0.0008]
(year==2011)*	-0.0062**	-0.0005	0.0015	-0.0013	-0.0069**	-0.0005
	[0.0030]	[0.0008]	[0.0048]	[0.0025]	[0.0031]	[0.0008]
(year==2012)*	-0.0048	-0.0005	0.0025	-0.0011	-0.0056*	-0.0005
	[0.0030]	[0.0008]	[0.0052]	[0.0027]	[0.0030]	[0.0008]
Year effects		Yes		Yes		Yes
Other controls		Yes		Yes		Yes
N		637		637		637

Notes: Observations are yearly and at the state level for the period 2000-2012. In each panel, the coefficients for the two sets of regressors (year phone adoption rate and year* rate of employment in Cable TV sector) are estimated from the same model, but reported next to each other for convenience. The estimates come from ordinary least squares regressions with state and year fixed effects. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE I.B-3: SECOND STAGE ESTIMATION USING THE TWO SETS OF INSTRUMENTS SEPARATELY

DV: Crime rate	(A) IV=Phone adoption rate			(B) IV=Rate of employment in Cable TV sector		
	Total	Violent	Property	Total	Violent	Property
Broadband lines rate	-0.0189**	-0.0049	-0.0207**	-0.0095*	-0.016	-0.0100*
	[0.0088]	[0.0122]	[0.0089]	[0.0050]	[0.0177]	[0.0053]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
N	636	636	636	656	656	656

Notes: Observations are yearly and at the state level for the period 2000-2012. The estimates come from 2SLS regressions with state and year fixed effects, whereby I use one set of IVs (year phone adoption rate in panel (A) and year* rate of employment in Cable TV sector in panel (B)) to instrument for broadband Internet penetration. Other controls include population density, unemployment rate, personal income per capita, percentage of Black or African American, percentage of males aged 15-24. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE I.B-4: FALSIFICATION TEST, 1990-2012

DV: Crime rate	(A) Total		(B) Violent		(C) Property	
Interaction	*Phone adoption rate	*Cable TV employment rate	*Phone adoption rate	*Cable TV employment rate	*Phone adoption rate	*Cable TV employment rate
(year==1991)*	-0.00114	0.00024	-0.00207	0.00121	-0.00094	0.00016
	[0.00103]	[0.00052]	[0.00172]	[0.00072]	[0.00103]	[0.00053]
(year==1992)*	-0.00197	-0.00021	-0.00503*	0.00197*	-0.00167	-0.00042
	[0.00159]	[0.00108]	[0.00268]	[0.00107]	[0.00147]	[0.00111]
(year==1993)*	-0.00324	-0.00069	-0.0046	0.00221	-0.0031	-0.00096
	[0.00213]	[0.00060]	[0.00327]	[0.00137]	[0.00198]	[0.00059]
(year==1994)*	-0.00632**	-0.00054	-0.00693*	0.0018	-0.00624**	-0.00074
	[0.00251]	[0.00125]	[0.00386]	[0.00186]	[0.00243]	[0.00125]
(year==1995)*	-0.00456	-0.00143	-0.00518	-0.00023	-0.00443	-0.00157
	[0.00276]	[0.00128]	[0.00443]	[0.00213]	[0.00278]	[0.00133]
(year==1996)*	-0.00751**	-0.00126	-0.00749	-0.00042	-0.00743**	-0.00139
	[0.00328]	[0.00163]	[0.00461]	[0.00207]	[0.00343]	[0.00173]
(year==1997)*	-0.01050***	-0.00144	-0.01094**	-0.00089	-0.01041***	-0.00156
	[0.00294]	[0.00186]	[0.00495]	[0.00230]	[0.00299]	[0.00193]
(year==1998)*	-0.01168***	-0.00167	-0.01410**	-0.00044	-0.01139***	-0.00186
	[0.00295]	[0.00134]	[0.00561]	[0.00202]	[0.00299]	[0.00139]
(year==1999)*	-0.01360***	-0.00139	-0.01669**	-0.0007	-0.01328***	-0.00157
	[0.00324]	[0.00143]	[0.00651]	[0.00230]	[0.00330]	[0.00149]
(year==2000)*	-0.01353***	-0.002	-0.01626***	-0.0025	-0.01328***	-0.00207
	[0.00314]	[0.00127]	[0.00583]	[0.00206]	[0.00322]	[0.00130]
(year==2001)*	-0.01310***	-0.00206	-0.01497***	-0.00192	-0.01289***	-0.00217
	[0.00345]	[0.00157]	[0.00526]	[0.00209]	[0.00364]	[0.00162]
(year==2002)*	-0.01353***	-0.00196	-0.01412***	-0.00217	-0.01341***	-0.00204
	[0.00382]	[0.00188]	[0.00513]	[0.00196]	[0.00413]	[0.00195]
(year==2003)*	-0.01469***	-0.00134	-0.01295**	-0.00219	-0.01497***	-0.00138
	[0.00381]	[0.00198]	[0.00555]	[0.00202]	[0.00400]	[0.00206]
(year==2004)*	-0.01800***	-0.00111	-0.01712***	-0.00119	-0.01823***	-0.00119
	[0.00367]	[0.00177]	[0.00635]	[0.00218]	[0.00372]	[0.00184]
(year==2005)*	-0.01812***	-0.00135	-0.01825**	-0.00218	-0.01849***	-0.00139
	[0.00406]	[0.00179]	[0.00750]	[0.00230]	[0.00407]	[0.00186]
(year==2006)*	-0.01728***	-0.00195	-0.02040**	-0.00365	-0.01736***	-0.00197
	[0.00420]	[0.00166]	[0.00826]	[0.00247]	[0.00425]	[0.00169]
(year==2007)*	-0.01820***	-0.00254	-0.02218**	-0.00424*	-0.01792***	-0.00258
	[0.00411]	[0.00168]	[0.00878]	[0.00237]	[0.00421]	[0.00168]
(year==2008)*	-0.01760***	-0.00241*	-0.02274**	-0.00335	-0.01718***	-0.00251*
	[0.00430]	[0.00143]	[0.00868]	[0.00243]	[0.00448]	[0.00143]
(year==2009)*	-0.01770***	-0.00256*	-0.02136**	-0.00373	-0.01745***	-0.00264*
	[0.00449]	[0.00145]	[0.00877]	[0.00247]	[0.00466]	[0.00143]
(year==2010)*	-0.01538***	-0.00229	-0.01810**	-0.00313	-0.01517***	-0.00239
	[0.00469]	[0.00157]	[0.00864]	[0.00270]	[0.00493]	[0.00155]
(year==2011)*	-0.01628***	-0.00248	-0.01845**	-0.00407	-0.01593***	-0.00254
	[0.00488]	[0.00159]	[0.00790]	[0.00274]	[0.00521]	[0.00160]
(year==2012)*	-0.01489***	-0.0026	-0.01527**	-0.00399	-0.01469***	-0.00266
	[0.00485]	[0.00161]	[0.00756]	[0.00299]	[0.00512]	[0.00159]
Year effects		Yes		Yes		Yes
Other controls		Yes		Yes		Yes
N		1,127		1,127		1,127

Notes: Observations are yearly and at the state level for the period 1990-2012. In each panel, the coefficients for the two sets of regressors (year phone adoption rate and year* rate of employment in Cable TV sector) are estimated from the same model, but reported next to each other for convenience. The estimates come from ordinary least squares regressions with state and year fixed effects. Other controls include population density, unemployment rate, personal income per capita, percentage of males aged 15-24. Robust standard errors (in brackets) are clustered at the state level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE I.B-5: FIRST STAGE ESTIMATION USING THE TWO SETS OF INSTRUMENTS SEPARATELY – MSA-LEVEL ANALYSIS

DV: HHs with Internet connection (%)		
Interaction	IV=Phone adoption rate	IV=Rate of employment in Cable TV sector
(year==2001)*	0.2018	-0.0224
	[5.0255]	[0.0351]
(year==2003)*	5.032	-0.0216
	[5.4283]	[0.0370]
(year==2007)*	-4.7255	0.025
	[4.2759]	[0.0415]
(year==2009)*	2.5751	0.0161
	[4.6101]	[0.0378]
(year==2010)*	-2.7061	-0.0083
	[5.4474]	[0.0374]
Year effects	Yes	Yes
Other controls	No	No
N	594	582
F-test for IV	1.38	0.51
F-test for IV - Prob > F	0.1806	0.7331

*Notes: Observations are yearly and at the MSA level for years 2000, 2001, 2003, 2007, 2009 and 2010. The estimates come from ordinary least squares regressions with MSA and year fixed effects. No other controls are included. "HH" stands for household. Robust standard errors (in brackets) are clustered at the MSA level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE I.B-6: SECOND STAGE ESTIMATION USING THE TWO SETS OF INSTRUMENTS SEPARATELY – MSA-LEVEL ANALYSIS

DV: Crime rate	(A) IV=Phone adoption rate			(B) IV=Rate of employment in Cable TV sector		
	Total	Violent	Property	Total	Violent	Property
HH has Internet (%)	-0.0084	0.0042	-0.0098	-0.0191	-0.0158	-0.0191
	[0.0149]	[0.0164]	[0.0149]	[0.0365]	[0.0385]	[0.0368]
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	No	No	No	No	No	No
N	584	584	584	569	569	569

Notes: Observations are yearly and at the MSA level for years 2000, 2001, 2003, 2007, 2009 and 2010. The estimates come from 2SLS regressions with state and year fixed effects, whereby I use one set of IVs (year phone adoption rate in panel (A) and year* rate of employment in Cable TV sector in panel (B)) to instrument for the percentage of households with an Internet connection. No other controls are included. "HH" stands for household. Robust standard errors (in brackets) are clustered at the MSA level. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

II.

Entertainment and crime: evidence from the 2014 World Cup in São Paulo

Abstract

This paper investigates the entertainment-crime relationship by analyzing high-frequency crime data for the city of São Paulo linked to information on the 2014 FIFA World Cup. Results show that crime is significantly lower during matches (-5.4%), especially those which likely envisage the highest remote viewership rates. I interpret these findings as reflecting voluntary incapacitation, whereby the potential for criminal interaction (and thus crime) drops as people are busy watching the games. I find no proof of short-run temporal displacement, suggesting that the crime reductions account for averted – rather than displaced – offenses. The main conclusion is that entertainment can play a crime-preventive role via voluntary incapacitation. However, this is not an unconditional statement. Results show that different outcomes generate heterogeneous effects in the games' aftermath: following Brazilian victories, crime increases significantly. Also, criminal activity reacts differently in the areas where game-related crowds assemble. There, voluntary incapacitation is more than offset by the positive effect of concentration.

Keywords: Entertainment • Crime • World Cup • Voluntary incapacitation.
JEL classification: K42.

II.1 Introduction

The idea that entertainment may exert a crime-preventing role is not a new one. The underlying rationale for this boils down to the old saying that “idle hands are the devil’s workshop”: as long as people are kept busy with some legal occupation that they enjoy, they will refrain from alternative activities that more likely lead to crime. This mechanism is known in the economics of crime literature as “voluntary incapacitation”. Crime prevention is the result of a time-use effect, whereby the potential for criminal interaction drops as potential offenders and victims voluntarily substitute time away from crime-conducive occupations to engage in alternative, less crime-conducive activities. This notion has even inspired some crime-preventing policies (for instance the Midnight Basketball initiative in the US), typically designed to encourage high-risk populations to participate in some subsidized recreational activity in order to divert them from more crime-conducive situations.

Despite the above considerations, the empirical evidence on the entertainment-crime link is quite scarce and scattered. In fact, there is no systematic set of contributions that target the effect of entertainment itself, but rather a handful of papers focusing on some specific aspects that are relevant to the relationship of interest. Results from this literature suggest that the entertainment-crime link is a complex one, which goes beyond the directness and simplicity of the “idle hands” – or voluntary incapacitation – logic. In particular, two dimensions seem to be relevant, namely concentration and content. An action that is performed in isolation (watching TV at home) is arguably different, as regards criminal outcomes, from one that entails an agglomeration (a big show). Then, the concentration dimension may play a role. In addition, entertainment does not only take up the audience’s time – it also transmits content, which may impact the audience’s propensity towards crime. On top of these considerations, there is the issue of temporal displacement. Whenever investigating the potential crime-reducing impact of keeping people busy (or otherwise incapacitated), it is crucial to assess whether crime is really being prevented or just displaced in time.

The present paper analyzes within one empirical exercise various aspects of the entertainment-crime relationship that have emerged from previous research, with the aim of providing a more unified set of results. To this end, I consider the entertainment event represented by the 2014 World Cup organized by the Fédération Internationale de Football

Association (FIFA). It is useful to make two preliminary observations that are relevant to the analysis. First of all, it is worth stressing that the audience of the event considered in this study mostly consists of remote (TV) viewers. Secondly, World Cup matches do not raise severe content-related worries, and are usually considered family-friendly entertainment. Instead, concerns are outcome-related, in that a match's result could affect the audience's propensity towards crime in the aftermath of the game. Bearing in mind these considerations, I exploit the tournament to study various aspects of the entertainment-crime relationship.

I use information on the 2014 FIFA World Cup linked to crime reports for the city of São Paulo. Observations are organized by the hour throughout the period 2006-2014. My strategy entails using high-frequency data to assess the impact of entertainment on crime, in a way that mimics an event study design. The use of high-frequency data is crucial to ensure that the estimated coefficients capture the impact of entertainment on crime, rather than that of some omitted factor. In fact, although there may be a number of unobserved variables that affect crime at the same time, none of them is likely to be discretely switched on and off according to exactly the same timing as the World Cup games. In addition, a rich set of time fixed effects is used to control for time-specific factors associated with crime. The source of exogenous variation in entertainment provision derives from games' scheduling: conditional on time fixed effects and further covariates, the scheduling of a game within a certain date and time is random with respect to crime (Laqueur and Copus, 2014).

I examine the overall entertainment-crime relationship by analyzing how criminal outcomes vary during the 2014 World Cup games, while controlling for a number of time fixed effects and other relevant covariates. Estimates indicate that the entertainment provided by World Cup matches leads to a 5.4% decrease in crime. I then further test whether this negative impact can be reasonably interpreted as reflecting a (dominant) voluntary incapacitation effect. The logic underlying the voluntary incapacitation notion implies that, the higher a game's audience, the stronger the voluntary incapacitation it generates. Since I do not have access to the TV viewership data for the city of São Paulo, I use some variables that are arguably correlated with that as proxies. In particular, I consider the games' kick-off time, worldwide live audience level and tournament round. Outcomes suggest that voluntary incapacitation is a sound explanation for the estimated negative impact of the matches on crime. In fact, significant decreases in criminal

activity are only found for games that are played in the evenings, have the highest worldwide live audience level and belong to the final rounds of the tournament.

I address the short-run temporal displacement hypothesis by checking the impact of the matches on the ±5 hours surrounding the event. I find no evidence of displacement, as crime does not appear to increase in the periods surrounding the games. On the contrary, it seems that the negative impact of the matches is extended to the hours right after and, possibly, right before the matches. This suggests that during these times potential offenders and victims engage in pre- and post-match activities, and this reduces the potential for criminal occurrences.

I investigate the outcome mechanism by considering whether criminal activity reacts differently to defeats of the Brazilian national team as compared to non-defeats (victories and draws) and further matches. Results show that indeed different outcomes produce heterogeneous effects in the games' aftermath. Following non-Brazilian games and Brazilian defeats, crime is significantly reduced. This is in line with the considerations made above about the apparent extension of the voluntary incapacitation effect to the hours surrounding the games. On the other hand, following a Brazilian non-defeat, crime significantly increases. This could be the result of concentration, as many supporters take the streets to celebrate the victory and create an environment that is more crime-conducive with respect to the alternative scenario (for instance, fewer people on the streets).

In the baseline analysis of the overall entertainment-crime relationship, I estimate that the coefficient associated with a game being played in São Paulo is positive, significant and substantial. A reasonable interpretation for this goes along the lines of the concentration effect: the agglomeration of large crowds around the stadium enhances the potential for criminal interaction during game times. To further investigate this feature, I convert my dataset into a panel by hour and police district ("PD") and check whether, in the stadium area and further neighborhoods that were preferred by the tourists to view the games, the evolution of criminal activity during World Cup matches follows a different pattern as compared to the rest of the city. Difference-in-differences estimates show that indeed these neighborhoods underwent a specific crime dynamic. While the simultaneous effect of matches on crime throughout the city is confirmed to be on average negative and significant (-6.6%), areas affected by game-related concentration register a positive and sizeable impact (59.5%).

I test the robustness of my results by replicating the analyses using an alternative set of time fixed effects, and find little cause for worry. Finally, I perform the estimations for specific offenses, using, as dependent variables, the number of reported robberies, thefts and assaults. Outcomes suggest that the overall negative impact of World Cup games on crime is mostly accounted for by the decrease in robberies (-9.5%). However, this conclusion should be taken with caution, as it may be driven to some extent by dissimilarities in the reporting patterns for different offenses – in particular, theft is often under-reported.

This work contributes to the existing scientific literature by investigating various aspects of the entertainment-crime relationship in a unified setting. In addition, my results speak to a highly policy-relevant debate by providing facts-based evidence on whether entertainment can soundly be exploited for crime prevention purposes. Finally, this paper offers novel insights by looking at the context of a developing country with high crime rates (Brazil ranks tenth in the World Bank list of countries by intentional homicide rate). As previous literature focuses on the developed world, this analysis represents an especially interesting ground for comparison.

The remainder of this paper is organized as follows. Section I.2 introduces the conceptual framework that motivates the relationship between entertainment and crime and discusses previous literature. In Section II.3 I present the data. Section II.4 explains the empirical strategy. Section II.5 describes my findings. Section II.6 discusses the results and concludes.

II.2 Conceptual framework and relevant literature

The idea that entertainment can be used to keep people busy and content so that they will refrain from alternative activities has deep roots. The word “diversion” itself derives from the Latin verb *divertere*, which literally means “turning aside from a course or purpose”. In ancient times, entertainment was exploited by the Roman emperors to keep the masses pleased and governable, according to the formula “bread and circuses” (*panem et circenses*). More recently, the hypothesis has emerged in criminology that entertainment may play a crime-preventive function (for instance, Zimring and Hawkins, 1997).

A widely-told story reckons that there was virtually no crime in Ney York City during The Beatles’ first appearance on US television, in 1964. This anecdote illustrates the appealing directness of crime prevention by diversion: entertainment averts crime by keeping potential

offenders and victims busy, thus reducing the potential for criminal interaction. Pursuant to this logic, the mechanism through which entertainment would affect crime is really that of incapacitation by a time-use effect. The literature mostly emphasizes the role of voluntary incapacitation on potential criminals. However, it is useful to remark that there is also a parallel effect on potential victims, who may be more or less susceptible to attacks while they engage in a specific activity with respect to the alternative occupation they would be pursuing otherwise. Although it is often hard to distinguish between the impacts on the two sides of criminal interaction, the interpretations of the results and the policy implications remain essentially unchanged.

The literature has recognized the importance of the voluntary incapacitation mechanism. Dahl and Della Vigna (2009) estimate that violent movie attendance reduces violent crime in the short term by between 1% and 2%, and interpret this finding along the lines of voluntary incapacitation. On a similar note, Cunningham, Engelstätter and Ward (2016) find that violent video games may reduce crime through the same mechanism.

In spite of the apparent directness and simplicity of the voluntary incapacitation effect, the relationship between entertainment and crime is more complex than this. In particular, two dimensions seem to be relevant. Firstly, the concentration dimension of a recreational activity may be relevant: an action that is performed in isolation (watching TV at home) is arguably different, as regards criminal outcomes, with respect to one that implies an agglomeration (a big show). Higher concentration is considered to be more crime conducive, as it increases the potential for interactions between potential victims and offenders and may reduce the chance of being caught (Kelly, 2000). Secondly, entertainment does not only take up the audience's time – it also transmits content, which may impact the audience's propensity towards crime. For instance, there is a wide-spread concern that violence and pornography in entertainment may promote criminal behavior by means of an arousal or an imitation effect. The content mechanism is not directly relevant to the entertainment event I consider in this study. In fact, World Cup games (like many other sporting events) are usually considered safe, family-friendly entertainment. An interesting feature here is rather that different game's results may generate heterogeneous effects on crime in the match's aftermath. Previous empirical contributions have acknowledged these critical aspects.

The importance of the concentration dimension of entertainment emerges from the literature that investigates the impact of sports-based entertainment events on crime. Big sporting events are likely to positively impact crime rates around the stadiums, mostly by means of a concentration effect. However, if the event is also broadcast on television, and television viewership is sufficiently large with respect to stadium attendance, it may still be that, at a less local level, the overall effect of the entertainment event on crime is negative – because the voluntary incapacitation of the television audience dominates the concentration effect.

Studies that focus on the effect of mass gatherings at sport stadiums typically find evidence that large entertainment events increase crime in the local community, due to concentration. Kalist and Lee (2016) show that home games of the of American National Football League ("NFL") increase total and property daily crime by 2.6%. The authors interpret their results as a consequence of concentration around the stadiums. Interestingly, they observe that crime decreases by almost 10% on playoff game days, just the days in which the ratio of television-watching audience to stadium crowds is likely to be at its highest. Rees and Schnepel (2009) study the relationship between college American football games and some specific offenses such as vandalism and assault using daily data. On game days, they find substantial crime increases in the host community. Marie (2016) uses information on stadium attendance at football games for major London teams combined with a panel of six-hourly crime data at the neighborhood level. He finds that property crime increases in the communities hosting the games. However, differently with respect to previous studies, he suggests that this result is driven by police displacement. Rees and Schnepel (2009), Kalist and Lee (2016) and Marie (2016) all focus on the effect of stadium attendance rather than remote viewership and, while showing that entertainment increases crime locally, they do not dismiss the possibility that, at a less local level, crime is actually lower because of the incapacitated remote audience.

Laqueur and Copus (2014) explicitly focus on big sports events for which stadium attendance is negligible as compared to television viewership. They analyze criminal activity in Chicago using half-hourly data and find substantial reductions in offenses during game times. This lends support to the hypothesis of a dominant crime-preventive role of entertainment via voluntary incapacitation, as long as the activity is not experienced in a mass concentration. Notice that Laqueur and Copus (2014)'s result is consistent with outcomes from Dahl and Della Vigna (2009) and Cunningham, Engelstätter and Ward (2016), both of which considered

entertainment activities that are typically experienced individually or in small groups (going to the movies and playing video games). Even the finding in Kalist and Lee (2016) that crime decreases on playoff game days points to the same conclusion. Indeed, playoff games are likely to display the highest ratio of television viewership to stadium attendance.

Kirby, Francis, and O'Flaherty (2014) study the association between viewing the FIFA World Cup remotely on television and domestic abuse in the United Kingdom. Using monthly and daily crime data referring to the 2002, 2006, and 2010 tournaments, the authors observe substantial increases on days in which the English national team played. The discrepancy between results from this study and Laqueur and Copus (2014) may be partially explained by the different time settings of the analyses. Using daily and monthly data, Kirby, Francis, and O'Flaherty (2014) cannot separate between the contemporaneous and the delayed effects of games. It may be that the former is negative (due to voluntary incapacitation) but more than offset by the latter. Other possible explanations are that entertainment may impact different offenses in a different way, and that results may vary depending on the context that is considered.

The outcome effect of sports-based entertainment has been addressed in the literature by exploring the link between game outcomes and crime. These studies typically stress the role of the emotional rationale, whereby different game's outcomes, and the discrepancy between expected and actual results, may make supporters frustrated or excited, with heterogeneous impacts on crime patterns. Along this line, Rees and Schnepel (2009) find that unexpected game outcomes lead to substantially larger increases in crime than non-upset outcomes. Card and Dahl (2011) observe that unexpected NFL game losses lead to a 10% increase in domestic violence, while neither expected losses nor upset wins have a significant effect. Finally, Kirby, Francis, and O'Flaherty (2013) find that the increases in domestic abuse are substantially higher on days in which the English national team lost.

On top of the outcome and concentration aspects, whenever one considers the crime incapacitation function of some activity or situation, it is crucial to address the temporal displacement issue. Imagine that over the restricted time window when a certain activity is ongoing – for instance an important game is broadcast on TV – the number of criminal incidents drops by means of a voluntary incapacitation effect. This occurrence does not necessarily imply that crime is being prevented. It may be that individuals simply displace all of their criminal activity to other times, so that the overall crime level remains unchanged over a broader time

window. In this case, the incapacitation function of entertainment would bring no benefit to society. Notice that the time displacement issue has implications as concerns the nature of criminal behavior. On the one hand, if by keeping people occupied delinquency is effectively prevented (i.e. there is no displacement), crime itself can be interpreted as opportunistic and situational: crime is averted by providing people something else to do. On the other hand, perfect crime displacement implies that offenders pre-determine the amount of crime they intend to perpetrate and organize their schedule accordingly.

Previous research has addressed the issue of short- or medium-term temporal displacement (while the empirical settings in this literature typically do not allow for testing long-run impacts). Most studies (for instance, Marie, 2016; Kalist and Lee, 2016; Laqueur and Copus, 2014; and Rees and Schnepel, 2009) conclude that there is no significant intertemporal crime substitution.

All in all, previous research shows that the relationship between entertainment and crime is a complex one. This study contributes to the existing literature by using novel data to analyze various aspects of the entertainment-crime relationship within one empirical exercise, with the aim of providing a more unified set of results.

II.3 Data

I investigate the entertainment-crime relationship using information on the 2014 FIFA World Cup combined with data from crime reports for the city of São Paulo, Brazil, in the period 2006-2014.

II.3.1 Information on the 2014 FIFA World Cup

The 2014 edition of the World Cup – the quadrennial world championship for men's national football teams organized by FIFA – took place in Brazil from June 12th 2014 to July 13th 2014, and ended with Germany’s victory. Thirty-two national teams played a total of 64 games in 12 venues located in as many host cities across Brazil, including São Paulo. There are a variety of publicly available statistics about the games. I display some of them in Table II.A-1. Six out of 64 games were played in São Paulo. The Brazilian team took part in 7 matches, and was defeated twice. All games were broadcast live and were scheduled to start at 13, 16, 17, 19 or 22 (São Paulo local time). The kick-off time was at 13 for most matches.

The regular duration of a match comprises two playing periods of 45 minutes plus a half-time break of 15 minutes. Throughout the analysis, I employ the approximation that a regular game lasts 2 hours. In all rounds of the tournament except for the initial one (the group stage), in case of a draw at the end of the normal playing time, there is an extension consisting of two extra 15-minute periods plus an interval of 5 minutes at the end of normal playing time (but not between the two extra periods). If the score is level after the extra time, penalty kicks are used to define the winner. Throughout the 2014 FIFA World Cup, a total of 8 games were determined through either extra time (4) or penalties (4). In these cases, I employ the approximation that a game lasts 3 hours.

At least one game was played on 25 out of the 32 days over which the tournament lasted, and most game days actually featured 3 matches, as displayed in Table II.A-2.

II.3.2 Data on crime

I obtained crime data from the Public Safety Department of the State of São Paulo (*Secretaria de Estado da Segurança Pública do Governo de São Paulo*). These entail crimes reported to the police that occurred in the city of São Paulo in the period 2006-2014. Data portray the criminal incidents in which the police completed a case report. They cover common offenses which I aggregated into six broader crime type categories, namely: murder (committed and attempted), robbery, rape, assault, theft and drugs-related offenses. All records provide information on the date of the occurrence as well as on the police district (*distrito policial*, “PD”) where it took place. I dropped the reports that do not indicate the time of the occurrence (about 22% of the total). There are 93 standard PDs in São Paulo. I aggregated the crime reports hourly both for the city as a whole and in panel (hourly by PD and 12-hourly by PD) setups. I use either one of these settings in my baseline analysis.

Panel (A) in Table II.A-3 reports some descriptive statistics on the total and specific offenses reported per hour in the city as a whole. It can be noticed that, as it is often the case with high-frequency criminal occurrences, the distributions of specific crimes are positively skewed. In addition, the high variances of most offenses suggest a situation of over-dispersion. Finally, as regards rapes, murders and drug-related crimes, the share of hours in which zero incidents are reported is very high.

Panel (B) in Table II.A-3 reports the corresponding descriptive statistics derived from the hourly/by PD panel. The crimes' distributions appear positively skewed, while there is no indication of over-dispersion in this setting. Notice that for 72% of hour/PD combinations, the total count of reported crimes equals zero. This figure is even higher for specific offenses.

In a robustness check, I use data organized in a 12-hourly panel (1 a.m. to midday or 1 p.m. to midnight) by PD. Panel (C) in Table II.A-3 reports the descriptive statistics derived from such setting. Again, skewness appears to be an issue. In addition, there is some evidence of over-dispersion as regards total crime and some specific offenses.

Panel (A) in Table II.A-4 reports the mean number of crimes per hour recorded in São Paulo as a whole by period of the day: 1 a.m. to midday or 1 p.m. to midnight. I choose these time windows to distinguish between the part of the day in which no World Cup game was played (1 a.m. to midday) from the part of the day in which all World Cup game were played (1 p.m. to midnight). Statistics are reported for different types of days (all days in the sample, non-game days, game days, days when a game was played in São Paulo, days when the Brazilian team played and did not lose, days when the Brazilian team played and was defeated).

Some interesting facts emerge. First of all, as expected, crimes are concentrated in the afternoon/evening period. More interestingly, while the crime level across day types remains rather stable in the night/morning hours, it varies considerably in the afternoon/evening period – that is when World Cup games are played. For instance, there is a 1.1 difference between the number of crimes per hour reported in the night/morning period on non-game *vs* game days, while the corresponding figure for the afternoon/evening period is 6.5. This is in line with the voluntary incapacitation hypothesis. Considering only game days in which a match was played in São Paulo, the difference between crimes per hour in the afternoon/evening period goes down to 3.8. This is in line with the hypothesis that, while the effect of games for the city as a whole is negative because of voluntary incapacitation, when games are played in São Paulo this negative impact is partially offset by increased criminal incidents around the points of game-related concentration.

I check whether there is some indication that local crime in areas where there is high fan concentration may react differently to the World Cup games with respect to the city as a whole. To this end, panels (B) and (C) in Table II.A-4 present the same statistics as panel (A) separating the specific PDs where fans assembled to watch the games. These include the stadium PD (panel

(B)) and two more areas – which I shall call “viewing areas” (panel (C)) – that the São Paulo Tourism Secretary identified as the tourists’ preferred locations to watch the games (the FIFA Fan Fest and Vila Madalena PDs). Remarkably, in these specific neighborhoods game days are associated with more crimes per hour in the afternoon/evening period than non-game days. This difference is even more extreme on days in which a game was played in São Paulo. On the other hand, the last panel of Table II.A-4 shows that on average other PDs (excluding the stadium and viewing ones) display the same patterns as the city as a whole (panel (A)). These statistics suggest that indeed the concentration effect is relevant to the entertainment-crime relationship.

II.4 Empirical strategy

To investigate the entertainment-crime relationship, I aggregate crime reports for the city of São Paulo by the hour, and link these data to information on football matches during the 2014 FIFA World Cup. The use of high-frequency data is crucial to ensure that the estimated coefficients capture the impact of entertainment on crime, rather than that of some omitted factor. In fact, although there may be a number of unobserved variables that affect crime at the same time, none of them is likely to be discretely switched on and off according to exactly the same timing as the World Cup games. Identification also relies on the assumption that, conditional on time fixed effects and further covariates, the scheduling of a game within a certain date and time is random with respect to crime. This seems reasonable and has been previously used in the literature (Laqueur and Copus, 2014). Notice that including a rich set of time fixed effects is crucial in order to ensure that the coefficients do not end up reflecting some omitted regularity in time together with the impact of interest.

II.4.1 Analysis of the hourly dataset

Considering the hourly dataset, the distribution of total crime appears over-dispersed (see panel (A) in Table II.A-3). Thus, I assume that this variable follows a negative binomial distribution, $Crime_t \backsim NegBi\ (\mu_t, \alpha)$, where $\mu_t = \exp(x_t^T \beta)$ and α is the dispersion parameter. Then, the distribution of $Crime_t$is:

(II.1) $$Pr(Crime_t) = \frac{\Gamma(Crime_t + \alpha^{-1})}{Crime_t! \Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_t}\right)^{\alpha^{-1}} \left(\frac{\mu_t}{\alpha^{-1} + \mu_t}\right)^{Crime_t}$$

Using the negative binomial specification with robust standard errors, the coefficients I estimate represent the effect of a 1-unit change in the respective independent variable on the logarithm of the expected incidence of the dependent variable ($Crime_t$), and can thus be interpreted as the percentage effect of the former on the latter. I estimate the impact of the entertainment event on criminal activity using the model described below:

(II.2) $$ln\, E\,(Crime_t) = \beta_0 + \beta_1 Game_t + \beta_2 BrGame_t + \beta_3 SPvenue_t + \beta_4 GameStatus_t + Z_t + \rho_t + \varepsilon_t$$

Where the subscript t refers to the hour-long time block; $Crime$ is the number of reported crimes; $Game$ is an indicator which equals 1 during 2014 World Cup games; $BrGame$ is a dummy for games in which the Brazilian national team participated; $SPvenue$ is an indicator for matches that were played in São Paulo; $GameStatus$ is a set of dummies including indicators for the hours right before and right after each game, as well as interactions indicating whether during the hour-long time block at least another World Cup game was played; ρ is a set of indicators for the time fixed effects that includes dummies for the year, month, day of the month, day of the week and hour. Finally, the vector Z comprises holiday indicators and variables that reflect changes occurred in the crime reporting and policing methods. In particular, I include a dummy which is equal to 1 starting from December 1st 2013, when it became possible to file reports for robbery online, and a dummy which is equal to 1 during the period May 20th 2014 to July 20th 2014, when a special police force was active to provide extra surveillance to the World Cup operations. The voluntary incapacitation effect is captured by variable $Game$, while $GameStatus$ indicators enable considerations on short-run temporal displacement.

I use as baseline dependent variable the total number of reported crimes per hour. In a few specifications, I replicate the analyses using as dependent variable the total number of reported robberies, assaults or thefts. I do not replicate the calculations for rapes, murders and drug-related offenses because for these variables the share of hours with zero reported incidents is very high, and this may impair the analysis.

I further test whether the negative simultaneous relationship between the games and crime can be reasonably interpreted as reflecting a (dominant) voluntary incapacitation effect. The logic underlying the voluntary incapacitation notion implies that, the higher a game's audience, the

stronger the voluntary incapacitation it generates. Since I do not have access to the TV viewership data for the city of São Paulo, I use some variables that are arguably correlated with that as proxies. In particular, I modify the model in Equation (II.2) to allow for heterogeneous effects depending on the games' characteristics. To this end, I construct three variations of Equation (II.2) by replacing $Game$ with different sets of mutually exclusive dummy variables that indicate the games' kick-off time, worldwide audience level and tournament round, respectively.

I investigate the outcome effect by allowing for heterogeneous effects depending on the result of games in which the Brazilian national team played. Brazil played 7 games in total, including 4 victories, 1 draw and 2 defeats. Starting from Equation (II.2), I substitute $Game$ with three mutually exclusive indicator variables, as displayed below:

(II.3) $$\ln \mathrm{E}(\boldsymbol{Crime_t}) = \boldsymbol{\gamma_0 + \gamma_1 NBrGame_t + \gamma_2 BrNDef_t + \gamma_3 BrDef_t + \gamma_4 SPvenue_t + \gamma_5 NBrGameStatus_t + \gamma_6 BrNDefStatus_t + \gamma_7 BrDefStatus_t + Z_t + \rho_t + \varepsilon_t}$$

Where $NBrGame$ is a dummy that equals 1 during World Cup games in which the Brazilian team did not participate; $BrNDef$ is an indicator for the World Cup games in which Brazil took part and did not lose (5 matches); and $BrDef$ an indicator for the World Cup games in which Brazil took part and was defeated (2 matches). Notice that I also differentiate the $GameStatus$ variables to allow for heterogeneous effects in the hours surrounding the games.

II.4.2 Analysis of the dataset by hour/police district

To study the effect of concentration, I organize data in a panel by hour and police district ("PD" – there are 93 standard PDs in São Paulo). Considering the hourly/by PD dataset, the distribution of total crime appears highly skewed, while over-dispersion does not appear to be an issue (see panel (B) in Table II.A-3). Thus, I assume that this variable follows a Poisson distribution, $Crime_{s,t} \backsim Poisson\ (\mu_{s,t})$, where $\mu_{s,t} = \exp(x_{s,t}^T \beta)$. Then, the distribution of $Crime_{s,t}$ is:

(II.4) $$Pr(Crime_{s,t}) = \frac{e^{-\mu_{s,t}} \mu_{s,t}{}^{Crime_{s,t}}}{Crime_{s,t}!}$$

Using the Poisson specification with fixed effects and robust standard errors clustered at the PD level, the coefficients I estimate represent the effect of a 1-unit change in the respective independent variable on the logarithm of the expected incidence of the dependent variable ($Crime_{s,t}$), and can thus be interpreted as the percentage effect of the former on the latter.

I check whether, in the neighborhoods which attract the biggest crowds during the World Cup matches, criminal activity follows a different pattern as compared to the rest of the city. I exploit the fact that 6 World Cup games were played in the São Paulo Arena – which was filled to capacity (about 63 thousands seats) on those occasions. I also use the information provided by the São Paulo Tourism Secretary that, apart from the stadium and family and friends' homes, the tourists' preferred locations to watch the games were the FIFA Fan Fest (a public viewing site organized by FIFA in the city center) and the Vila Madalena neighborhood. I perform a difference-in-differences ("DID") analysis of the effect of concentration within the Poisson setting using the following model:

(II.5) $$ln\, E\,(Crime_{s,t}) = \delta_0 + \delta_1[(SPvenue_t * StadiumPD_s) + (Game_t * ViewingPD_s)] + \delta_2 Game_t + \delta_3 BrGame_t + \delta_4 SPvenue_t + Z_t + \rho_t + \theta_s + \varepsilon_{s,t}$$

Where the subscripts s and t denote the PD and the hour-long time block respectively; $StadiumPD$ is a dummy for the PD where the stadium is located; $ViewingPD$ is an indicator for the PDs where the FIFA Fan Fest or Vila Madalena are located; and θ are PD fixed effects. In this specification, coefficient δ_1 portrays the differential impact of games on crime in the areas subject to game-related concentration, while coefficient δ_2 captures the average effect on the city's PDs.

A concern with this setup is that standard inference methods used in DID models may not perform well because there are only 3 treated PDs. In particular, Conley and Taber (2011) show that if the number of treated units (N_1) and the number of periods are fixed, the DID estimator is unbiased but inconsistent, so that the estimated treatment coefficient tends in probability to the true parameter of interest plus a noise ($\widehat{\delta_1} \rightarrow \delta_1 + W$). Then, using standard inference methods would result in misleading standard errors. The authors develop an alternative approach to inference under the assumption that N_1 is small. This approach allows for the calculation of

reliable confidence intervals for the treatment coefficient. To alleviate concerns stemming from the small number of treatment units, I replicate the analysis using Conley and Taber's inference method to calculate the confidence intervals for the coefficient of interest (δ_1).

II.5 Results

II.5.1 Main results from the analysis of the hourly dataset

II.5.1.1 The impact of World Cup games on crime in São Paulo

Table II.A-5 presents the negative binomial regression estimates from Equation (II.2). Column (1) displays results from a simple specification which only considers the simultaneous effect of the matches on crime ($\beta_4 = 0$). In specifications (2) and (3), I sequentially add $GameStatus$ dummies for, respectively, 2 and 5 hours before and after the games, as well as the interactions indicating whether at least another World Cup game is being played at the same time. Results considering 1, 3 and 4 hours before and after the games are also available upon request. Let us first focus on specification (1). The voluntary incapacitation effect, captured by the coefficient for $Game$, is negative and significant. Estimates indicate that the entertainment provided by World Cup games leads to a 5.4% decrease in crime. Whether or not the Brazilian national team is one of the players does not seem to make a difference. The coefficient associated with the game being played in São Paulo is positive, significant and more than twice as large (in absolute value) than the estimated voluntary incapacitation effect. Similar considerations hold for specification (2), while the estimated coefficients lose significance in specification (3). This is likely caused by the fact that most tournament days feature more than one match, played with a lag of 3-4 hours from each other. Thus, collinearity between $Game$ and some $GameStatus$ indicators may be an issue. As will be shown below, allowing for heterogeneity of the effects across different kinds of games (for instance, by kick-off times) leads to more precisely estimated results.

What emerges clearly from specifications (2) and (3) is the lack of evidence of short-run temporal displacement. On the contrary, it seems that the negative impact of the games on crime is extended to the subsequent period and, possibly (specification (2)), to the hour just before the

match. These results suggest not only that potential offenders do not displace crime to the periods surrounding the games, but also that during these times they engage in pre- and post-match activities, and are thus incapacitated. In particular, in the second to fourth hours after the games, crime is significantly reduced by 10.1% to 17.8% – a proportion that is even more substantial than the contemporaneous effect. This outcome appears rather extreme, and is likely caused by the collinearity issue mentioned above. Allowing for heterogeneity of the effects across different kinds of games addresses this concern.

As a final remark, the estimated dispersion parameter α indicates that over-dispersion is indeed an issue, therefore it is appropriate to use the negative binomial model.

I further test whether the negative simultaneous relationship between games and crime can be reasonably interpreted as reflecting a (dominant) voluntary incapacitation impact. In particular, I allow for heterogeneous effects depending on the games' kick-off time, worldwide audience level and tournament round. I use these variables as proxies for the games' TV viewership level in the city of São Paulo. Results are displayed in Table II.A-6. Specifications (1) to (3) follow the same logic as in Table II.A-5. It appears from panel (A) in Table II.A-6 that the only games capable of generating a significant impact are those broadcast in the evening. In other words, only in the evening hours a significant amount of crimes is averted because people are busy watching the game. This appears reasonable because (i) in general, most crimes are committed in the evening hours; and (ii) evening games arguably feature the highest audience levels. Both of these occurrences are ultimately due to the fact that the evening hours are those when people are most likely to be idle, creating the highest potential for crimes to be committed or averted. The simultaneous effect of evening games on crime is significant and substantial, with estimated coefficients that range from -10.8% to -25.4%. Whether or not the game is played by the Brazilian national team or in São Paulo does not seem to affect this outcome. There is no evidence of short-run temporal displacement. On the contrary, the entertainment provided by World Cup matches seems to reduce crime also in the hours surrounding the games. In particular, in the second to fourth hours after the game, crime significantly decreases by a proportion that, in each specification, is close to – but smaller in size than – the contemporaneous effect.

Panel (B) in Table II.A-6 displays evidence of additional heterogeneity in the games-crime relationship. A significant impact is only found for games with the maximum worldwide live viewership level. This appears reasonable, as higher audience implies more incapacitated

individuals. Notice that FIFA only reports audience levels worldwide, while I am really interested in the São Paulo figures. However, given that the latter are unavailable, it is reasonable to assume that viewership patterns in São Paulo are in line with the global ones. During matches with an average live audience of 250 million people or more, crime decreases by 18.9% to 21.7%, depending on the specification. This outcome is partially reversed in case the Brazilian national team takes part in the match, or in case the game is played in São Paulo. Again, results confirm that the negative relationship extends to the hours surrounding the game.

Panel (C) in Table II.A-6 shows that the impact of World Cup games on crime also depends on the tournament stage. The initial stages have an ambiguous or even significantly positive impact on crime, whereas starting from the round-of-8 games (quarter-finals) the effect is negative, significant and substantial, ranging from -23.3% to -34.5%. This seems reasonable, as final-stages games are likely to attract a larger audience and thus deploy a stronger voluntary incapacitation impact. This outcome is not significantly affected in case the Brazilian national team takes part in the game, while there is some evidence that a match being played in São Paulo positively affects criminal activity. Again, it is confirmed that offenses are not temporally displaced to the hours surrounding the matches. On the contrary, the negative impact seems to extend to these hours.

In conclusion, these heterogeneity exercises increase my confidence that the negative games-crime relationship reflects the voluntary incapacitation effect.

II.5.1.2 The impact of games' outcome

Previous literature has shown that the impact of a game on crime in the hours following a match may differ depending on the game's outcome. Studies have typically focused on the emotional rationale: supporters can get upset or excited about a game's result, and this may impact criminal outcomes (for instance, Card and Dahl, 2011). However, this is not the only possible driver for heterogeneous outcomes. For instance, different results may affect the concentration patterns in the matches' aftermath, thus impacting the potential for criminal interaction. In particular, it is often the case after important victories that crowds of supporters take the streets to celebrate, creating a crime-prone environment.

I test whether different games' outcomes generate heterogeneous effects in my sample. Table II.A-7 reports the negative binomial regression estimates from Equation (II.3), where I

allow for separate effects of Brazil's defeats (2 games) and non-defeats (5 games – 4 victories and 1 draw). Specifications (1) to (3) follow the same logic as in Table II.A-5. Results portray a very different pattern in the aftermath of Brazilian non-defeats with respect to both Brazilian losses and games in which the Brazilian squad does not participate. In particular, in the hours that follow a Brazilian victory (non-defeat), crime increases significantly and substantially, with estimated peaks of up to 26.3%. This could be the result of concentration, as many supporters take the streets to celebrate and create an environment that is more conducive to some kinds of crime (for instance alcohol-related offenses or street fights) with respect to the alternative scenario (for instance, fewer people on the streets). On the other hand, in the hours following non-Brazil games crime is significantly reduced, in line with baseline findings in Table II.A-5. Also the estimated coefficients for the hours following a Brazilian loss are negative and significant, suggesting that after a defeat fans may be less willing to party and opt for less crime-conducive situations (for instance, they may go – or stay – at home). Notice that, concerning the hours before the games, a significant and negative effect is only found for Brazilian defeats. This is puzzling, as a match's outcome can hardly explain the pre-game dynamics. Reasonably, results regarding the hours surrounding Brazilian losses may not be reliable as they are based on just two games.

Finally, the estimated simultaneous impact on criminal activity of matches in which the Brazilian team took part is negative (for both defeats and non-defeats), but not precisely estimated. This is surprising, as one would expect national team's games to generate the strongest voluntary incapacitation effect in São Paulo.

II.5.2 Further results from the analysis of the hourly dataset

II.5.2.1 Previous tournament editions

The FIFA World Cup is played every four year. Therefore, two further tournament editions (besides the 2014 one) were played in the sample period: one in 2006 and one in 2010. The 2006 competition was held from June 9th to July 9th in Germany, and ended with Italy's victory. In 2006, Brazil played a total of 5 games (4 victories and 1 defeat), and was eliminated by France in the quarter-finals. The 2010 World Cup took place in South Africa from June 11th to July 11th and

was won by Spain. Also in 2010, Brazil played a total of 5 games (3 victories, 1 draw and 1 defeat), and was knocked out by the Netherlands in the quarter-finals.

It is interesting to check how crime in São Paulo behaved during the Brazilian team's matches in previous World Cup editions, and how this compares to the pattern during the 2014 games. To this end, I modify Equation (II.2) to include an indicator that is equal to 1 during Brazil's games in previous editions of the tournament. I expect viewership in São Paulo to be high enough to generate some sizeable impact only for games in which the national squad took part. This is due to two reasons. First of all, because of the time difference between Brazil and the tournament venues, many games were broadcast at inconvenient times for the Brazilian audience. Secondly, the enthusiasm and interest in the event itself (apart from the national squads' matches) were arguably lower for World Cup editions that did not take place in Brazil.

Results are displayed in Table II.A-8. During Brazil's matches in previous World Cups, crime in São Paulo significantly decreased by 11.6%. This suggests, as expected, that the voluntary incapacitation mechanism was at play during these games. The effect is about twice as large (in absolute value) as that estimated for 2014 games. This difference in size may represent the impact of concentration. Concentration (during games) in São Paulo was arguably much higher in 2014 than in previous tournaments. Then, the previous World Cup games' coefficient would be "more purely" reflecting the voluntary incapacitation mechanism – while the 2014 games' coefficient also captures some share of the concentration impact. This result further stresses the importance of taking into account the concentration dimension when evaluating the impact of an entertainment event on crime.

II.5.2.2 Robustness

A crucial element of my identification strategy is that time fixed effects are properly controlled for, so that the coefficients do not reflect some omitted regularity in time. For instance, crime is higher in the afternoon/evening than in the night/morning, as displayed in Table II.A-4. All tournament games took place in the afternoon/evening hours. Then, failing to properly account for hour fixed effects would lead to an omitted variable bias, as the estimated game's coefficient would reflect the hour fixed effect as well as the impact of interest.

In all calculations I include indicators for the year, month, day of the month, day of the week and hour. In this robustness test, I check whether including day-of-the-year effects affects my results. I estimate for each of the analyses the specification which only considers the simultaneous effect of the matches on crime (specification (1)) including dummies for the year, day of year (dropping February 29th in leap years), day of the week and hour. Panel (A) in Table II.A-9 shows the results. Outcomes are robust to this modification of included time fixed effects as precisely estimated coefficients are in line with those obtained in the baseline specifications.

II.5.2.3 Specific crimes

In order to get a deeper understanding of the effect on crime of the entertainment generated by World Cup games, I replicate each of the analyses using, as dependent variables, the total number of reported robberies, thefts and assaults. I each case, I estimate the specification which only considers the simultaneous effect of the matches on crime (specification (1)).

Results for robberies, thefts and assaults are reported in Table II.A-9 in panels (B), (C) and (D), respectively. Robberies decrease by 9.5% during game times, and envisage analogous heterogeneities as those found for total crime in the specifications by kick-off time, audience and tournament round. On the other hand, the estimated impacts on thefts and assaults are seldom significant. As regards the latter, a positive and significant (at 10%) simultaneous effect is found for games in which the Brazilian national team won. This outcome is consistent with the hypothesis that Brazil’s victories attract people to the streets to celebrate, generating a concentration effect that increases the potential for some kinds of crime (such as street fights). All in all, robbery seems to be the offense that is most significantly impacted by the World Cup games. However, this conclusion should be taken with caution, as it may be driven to some extent by dissimilarities in the reporting patterns for different offenses. In particular, it is often the case that theft is under-reported.

II.5.3 Main results from the analysis of the dataset by hour/police district: concentration

An entertainment activity that is experienced in isolation may be different, as regards criminal outcomes, with respect to one that implies concentration. In particular, higher

concentration is considered to be more crime-conducive, as it increases the potential for interactions between potential victims and offenders and may reduce the chance of being caught (Kelly, 2000). In the baseline analysis of the overall entertainment-crime relationship, I estimate that the coefficient associated with a game being played in São Paulo is positive, significant and more than twice as large (in absolute value) as the voluntary incapacitation effect. I use the data organized in a panel by hour and PD to further investigate concentration. In particular, I test whether, in the neighborhoods which attracted the biggest crowds during the World Cup matches, criminal activity follows a different pattern as compared to the rest of the city. The São Paulo Tourism Secretary identified the viewing areas (the FIFA Fan Fest and Vila Madalena PDs) as the tourists' preferred locations to watch all games. In addition, the stadium PD attracted large numbers of fans to watch the 6 matches which were played in the São Paulo Arena. Thus, I consider that the PDs where fans concentrated the most to watch the competition include the viewing areas (for all games) and the stadium PD (only for games played in São Paulo).

Column (A) in Table II.A-10 presents estimated coefficients from Equation (II.5). Notice that these derive from a Poisson regression, as the relevant descriptive statistics show no sign of over-dispersion (see panel (B) in Table II.A-3). Anyway, outcomes remain essentially unchanged using the negative binomial specification. Results lend support to the hypothesis that the concentration effect dominates the voluntary incapacitation one in those areas that attract the biggest crowds during World Cup games. While the baseline result is confirmed that match times are associated with a 6.6% decrease in crime on average in São Paulo's neighborhoods, in the areas where fans assemble to view the matches this negative voluntary incapacitation effect is more than outweighed by a positive and substantial (59.5%) concentration effect.

As reported in the second panel of Table II.A-3, for 72% of hour/PD combinations, the total count of reported crimes equals zero. This is not ideal for my analysis, as the model may not fit the data properly. To alleviate this worry, I replicate the analysis at the 12-hourly/PD level in section 5.4.1. A further concern with this analysis stems from the fact that there are only 3 treated PDs. Therefore, using the standard inference methods could result in misleading standard errors. To address this concern, I replicate the analysis using Conley and Taber (2011)'s inference method to calculate the confidence intervals for the concentration effect's coefficient. I compare 95% confidence intervals from this and the standard (cluster-robust) approach in Table II.A-11. Outcomes confirm that the effect of concentration on crime is significant.

II.5.4 Further results on concentration

II.5.4.1 Robustness

A major concern with the concentration analysis is that the model may not properly fit the data because of the high number of zero-count observations in the panel by hour/PD. In order to alleviate this concern, I collapse the time dimension of the dataset and convert it to a 12-hourly panel by PD. I divide each day into 2 time periods: 1 a.m. to midday and 1 p.m. to midnight. I choose these time windows to distinguish between the part of the day in which no World Cup game was played (1 a.m. to midday) from the part of the day in which all World Cup game were played (1 p.m. to midnight). This allows me to replicate the concentration analysis while avoiding the excessive zero-count observations problem for the total crime variable, as displayed in panel (C) in Table II.A-3. The shortcoming of this specification is that it is more likely to suffer from an omitted variable bias, as I am renouncing – to a certain extent – to the high-frequency feature of the data. However, it is useful, as a robustness test, to replicate the concentration analysis and compare the results.

Since there is some indication of over-dispersion as regards the total crime variable distribution (see panel (C) in Table II.A-3), following Marie (2016) I employ the negative binomial model with fixed effects. Column (B) in Table II.A-10 shows the results from this robustness check. The estimated coefficients for the simultaneous effects of the World Cup games on crime are bigger in absolute value but consistent with those from the analysis by hour/PD in terms of both sign and significance. This enhances confidence that the high number of zero-count observations does not invalidate the baseline concentration analysis.

Within the above setting (12-hourly by PD) I replicate the analysis using the alternative time fixed effects, which include dummies for the year, day of year (dropping February 29th in leap years), day of the week and 12-hour period. I do not perform this estimation using the hour/PD panel because it is computationally too demanding. Column (C) in Table II.A-10 shows that this modification of included time fixed effects has little impact on the results. Although I cannot perform this check on my preferred panel setting (by hour/PD), this outcome indirectly provides some indication that the results are robust to the alternative time fixed effects specification.

II.5.4.2 Specific crimes

It is possible to replicate the analysis for some specific offenses in the 12-hourly/PD panel setting, as the excessive zero-count observations problem is alleviated (see panel (C) in Table II.A-3). Column (D) and (E) in Table II.A-10 report the results for robberies and thefts, respectively, using the baseline time fixed effects (dummies for the year, month, day of the month, day of the week and 12-hour period). Results show that, in the neighborhoods preferred by the tourists to watch the competition, both robberies and thefts increase significantly during the games. The coefficient for thefts is especially high (92.9%), but not unreasonable. It is easy to picture that, while victims are focused on watching the games in some very crowded location, offenders have an easy time pick-pocketing their wallet or mobile phone. For both robberies and thefts, the effect of World Cup games on São Paulo's neighborhoods is on average negative and significant, suggesting that the voluntary incapacitation mechanism is at play.

II.6 Discussion

In line with Laquer and Copus (2014), Dahl and Della Vigna (2009), and Cunningham, Engelstätter and Ward (2016) my outcomes confirm that entertainment can play a crime-preventive role by voluntary incapacitation. I estimate that during World Cup games, total crime per hour in São Paulo decreased significantly by 5.4%. To put this result into perspective, it is useful to perform a back-of-the-envelope calculation. On non-game days an average of 42.42 crimes per hour are reported across the city in the afternoon/evening period (1 p.m. to midnight). Then, the voluntary incapacitation generated by a World Cup game leads to 2.3 fewer crimes per hour. During the 2014 tournament, games extended over a total of 120 hours, all in the afternoon/evening period. Then, World Cup-related voluntary incapacitation averted 275 offenses in total.

In addition, my outcomes suggest that the reductions in crime really account for prevented – rather than temporally displaced – offenses, in agreement with the vast majority of papers in the cited literature. This in turn provides support to the hypothesis that some portion of crime is opportunistic and situational – it can be averted by keeping people busy and content with something else.

My findings suggest that the outcome-related dimension is important, as different game results generate opposite consequences in terms of criminal activity. The likely rationale underlying this heterogeneity of effects is represented by the concentration mechanism, whereby crowds of supporters take the streets after the national team's victories, creating a more crime-prone environment. By suggesting that concentration may be the driver for heterogeneous post-game effects, my findings stand apart from previous literature on the topic (for instance, Card and Dahl, 2011), which typically stresses the emotional rationale.

The importance of the concentration dimension also emerges from the finding that, while World Cup matches generate a crime reduction via voluntary incapacitation throughout the city as a whole, this impact is more than offset in areas affected by game-related agglomeration. This result is in line with conclusions from previous studies which focus on the effect of mass gatherings at sport stadiums, such as Rees and Schnepel (2009) and Kalist and Lee (2016). To better interpret my findings, it is useful to perform a back-of-the-envelope calculation to compare the relative importance of the concentration and incapacitation effect. The latter mechanism accounts on average for a 6.6% decrease in crime in each PD. On top of this, I estimate that concentration generates a 59.5% increase in crime in the two viewing PDs and, during games played in São Paulo, in the stadium PD. Taking as baseline the average number of crimes per hour in the different PDs in the afternoon/evening period on non-game days (as reported in Table II.A-4), it appears that, in line with results from the city-level analysis, the voluntary incapacitation effect leads to 2.8 fewer crimes per hour. Once the positive impact of concentration in the viewing and stadium PDs is factored in, the crime drop decreases to 1.8 fewer occurrences per hour during games that are not played in São Paulo and 1.7 fewer crimes per hour during games that are played in São Paulo. This implies that, although voluntary incapacitation dominates concentration, the latter is extremely important and capable of undercutting the impact of the former.

Speaking to the policy-relevant debate on crime prevention, this study finds evidence that entertainment can play a crime-preventive role via voluntary incapacitation, and could thus be exploited as policy tool. Clearly, this should be done with caution, paying special attention to the outcome and concentration dimensions, as well as the specific context of application.

Appendix II.A Tables

TABLE II.A-1: WORLD CUP GAMES' SUMMARY STATISTICS

	Number of games	Share of games
All games	64	100%
Games in São Paulo	6	9%
Time extension/penalties	8	13%
Brazil games	7	11%
Wins	4	6%
Draws	1	2%
Defeats	2	3%
Games by kick-off time		
13	24	38%
16	11	17%
17	18	28%
19	10	16%
22	1	2%
Games by worldwide live audience level		
> 250 mln	13	20%
< 250 mln	51	80%
Games by tournament round		
Group stage	48	75%
Round of 16	8	13%
Quarter finals	4	6%
Finals and semi-finals	4	6%

Notes: Summary statistics for the 64 games played during the 2014 FIFA World Cup.

TABLE II.A-2: TOURNAMENT DAYS FEATURING *N* GAMES

n	**Number of days**	**Share of days**
0	7	22%
1	5	16%
2	6	19%
3	9	28%
4	5	16%
Total	32	100%

Notes: Summary statistics for the 32 days over which the 2014 FIFA World Cup was played (June 12th 2014 to July 13th 2014).

TABLE II.A-3: CRIMES PER HOUR, 2006-2014 – DESCRIPTIVE STATISTICS

	Total crime	Robbery	Assault	Rape	Murder	Theft	Drugs
A) Dataset by hour							
Mean	33.45	16.20	3.98	0.15	0.31	12.19	0.62
Min	0	0	0	0	0	0	0
Max	122	85	31	5	7	58	31
Obs=0 (%)	0.01	0.24	8	86.65	74.53	0.86	59.49
Variance	271.17	106.23	8.59	0.16	0.35	55.09	0.90
Skewness	0.37	1.08	1.07	2.83	2.16	0.47	2.62
N	78,840	78,840	78,840	78,840	78,840	78,840	78,840
B) Dataset by hour/PD							
Mean	0.36	0.17	0.04	0.00	0.00	0.13	0.01
Min	0	0	0	0	0	0	0
Max	34	23	7	2	3	28	7
Obs=0 (%)	72.03	84.90	95.89	99.86	99.67	88.46	99.35
Variance	0.44	0.20	0.04	0.00	0.00	0.15	0.01
Skewness	2.38	2.94	5.11	25.50	17.81	3.86	13.10
N	7,332,120	7,332,120	7,332,120	7,332,120	7,332,120	7,332,120	7,332,120
C) Dataset by 12 hours/PD							
Mean	4.32	2.09	0.51	0.02	0.04	1.57	0.08
Min	0	0	0	0	0	0	0
Max	148	91	10	3	3	118	9
Obs=0 (%)	5.78	22.02	63.1	98.14	96.18	30.88	92.77
Variance	10.94	4.03	0.64	0.02	0.04	3.44	0.09
Skewness	2.12	1.66	1.91	7.49	5.38	3.78	4.38
N	611,010	611,010	611,010	611,010	611,010	611,010	611,010

Notes: "Murder" comprises committed and attempted murders. In panel (A) summary statistics are generated from the crimes reported hourly in the city of São Paulo as a whole over the period 2006-2014. In panel (B) summary statistics are generated from the crimes reported hourly in each of the city's 93 PDs over the period 2006-2014. In panel (C) summary statistics are generated from the crimes reported 12-hourly (1 a.m. to midday and 1 p.m. to midnight) in each of the city's 93 PDs over the period 2006-2014.

TABLE II.A-4: NUMBER OF CRIMES PER HOUR BY PERIOD OF THE DAY AND TYPE OF DAY, 2006-2014

Day type	Mean (Standard deviation) 1 a.m. to Midday	Mean (Standard deviation) 1 p.m. to Midnight	N	Mean (Standard deviation) 1 a.m. to Midday	Mean (Standard deviation) 1 p.m. to Midnight	N
	A) Dataset by hour – São Paulo			**B) Dataset by hour/PD – Stadium PD**		
All	24.54	42.37	78,840	0.15	0.28	78,840
	(9.77)	(9.48)		(0.06)	(0.07)	
No game	24.54	42.42	78,240	0.15	0.28	78,240
	(9.79)	(9.50)		(0.06)	(0.07)	
Game	23.46	35.92	600	0.24	0.66	600
	(7.74)	(7.91)		(0.33)	(0.41)	
Game in São Paulo	23.43	38.63	144	0.64	2.04	144
	(8.73)	(9.39)		(1.31)	(1.82)	
Brazil wins/draw	24.03	42.20	120	0.38	1.07	120
	(9.16)	(11.92)		(0.64)	(0.79)	
Brazil defeats	22.08	33.50	48	0.13	0.29	48
	(10.39)	(9.42)		(0.22)	(0.44)	
	C) Dataset by hour/PD – Viewing PDs			**D) Dataset by hour/PD – All but stadium and viewing PDs**		
All	0.5	0.8	157,680	0.26	0.45	7,095,600
	(0.16)	(0.12)		(0.1)	(0.1)	
No game	0.5	0.8	156,480	0.26	0.45	7,041,600
	(0.17)	(0.12)		(0.1)	(0.1)	
Game	0.45	1.1	1,200	0.25	0.37	54,000
	(0.17)	(0)		(0.08)	(0.08)	
Game in São Paulo	0.41	1.27	288	0.24	0.38	12,960
	(0.24)	(0.41)		(0.08)	(0.1)	
Brazil wins/draw	0.51	2.18	240	0.25	0.41	10,800
	(0.27)	(0.76)		(0.09)	(0.12)	
Brazil defeats	0.35	1.42	96	0.24	0.34	4,320
	(0.38)	(0.83)		(0.11)	(0.09)	

Notes: In panel (A) 12-hourly summary statistics (1 a.m. to midday or 1 p.m. to midnight) are generated from the crimes reported hourly in the city of São Paulo as a whole over the period 2006-2014. In panel (B) 12-hourly summary statistics (1 a.m. to midday or 1 p.m. to midnight) are generated from the crimes reported hourly in the stadium PD over the period 2006-2014. In panel (C) 12-hourly summary statistics (1 a.m. to midday or 1 p.m. to midnight) are generated from the crimes reported hourly in the viewing PDs over the period 2006-2014. In panel (D) 12-hourly summary statistics (1 a.m. to midday or 1 p.m. to midnight) are generated from the crimes reported hourly in each of the city's PDs excluding the stadium and viewing ones over the period 2006-2014.

TABLE II.A-5: THE IMPACT OF WORLD CUP GAMES ON CRIME IN SÃO PAULO

DV: Total crime	(1)	(2)	(3)
Game	-0.054**	-0.066*	-0.057
	[0.026]	[0.034]	[0.078]
Brazil game	0.007	0.008	-0.001
	[0.067]	[0.067]	[0.076]
Game in São Paulo	0.143*	0.141*	0.114
	[0.080]	[0.081]	[0.080]
1 Hour(s) after game		-0.025	0.013
		[0.033]	[0.039]
2 Hour(s) after game		-0.101**	-0.114**
		[0.044]	[0.047]
3 Hour(s) after game			-0.178***
			[0.051]
4 Hour(s) after game			-0.119**
			[0.049]
5 Hour(s) after game			-0.013
			[0.044]
1 Hour(s) before game		-0.055*	-0.023
		[0.033]	[0.040]
2 Hour(s) before game		-0.058	-0.064
		[0.040]	[0.043]
3 Hour(s) before game			0.021
			[0.056]
4 Hour(s) before game			-0.061
			[0.044]
5 Hour(s) before game			-0.043
			[0.041]
lnalpha	-3.163***	-3.163***	-3.164***
	[0.012]	[0.012]	[0.012]
Time fixed effects	Yes	Yes	Yes
Status interactions	No	Yes	Yes
Other controls (Z_t)	Yes	Yes	Yes
N	78,840	78,840	78,840

*Notes: An observation is an hour-long period for the city of São Paulo as a whole over the period 2006-2014. The estimates come from negative binomial regressions; robust standard errors are in brackets. "Lnalpha" is the estimate of the log of the dispersion parameter. Time fixed effects include dummies for the year, month, day of the month, day of the week and hour. Status interactions indicate whether more than one World Cup games are played during the hour-long time block. Other controls (Z_t) include holiday indicators and variables that reflect changes occurred in the crime reporting and policing methods. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE II.A-6: THE IMPACT OF WORLD CUP GAMES ON CRIME IN SÃO PAULO – HETEROGENEOUS EFFECTS

DV: Total crime	**A) By kick-off time**			**B) By worldwide audience level**			**C) By tournament round**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Game starts at 13hs	-0.018	-0.027	0.074						
	[0.041]	[0.044]	[0.089]						
Game starts at 16hs or 17hs	-0.055	-0.075	-0.107						
	[0.037]	[0.046]	[0.087]						
Game starts at 19hs or 22hs	-0.108**	-0.139***	-0.254**						
	[0.044]	[0.049]	[0.122]						
WW avg live audience >250 mln				-0.209***	-0.217***	-0.189**			
				[0.064]	[0.062]	[0.080]			
WW avg live audience <250 mln				-0.034	-0.041	0.053			
				[0.027]	[0.035]	[0.085]			
Group-stage game							-0.034	-0.019	0.224**
							[0.027]	[0.037]	[0.095]
Round-of-16 game							0.016	0.02	0.203**
							[0.050]	[0.052]	[0.087]
Round-of-8 game							-0.339***	-0.345***	-0.137
							[0.103]	[0.103]	[0.129]
Round-of-4 game							-0.259***	-0.264***	-0.233***
							[0.082]	[0.082]	[0.079]
Brazil game	0.003	0.008	0.008	0.135*	0.134*	0.143*	0.072	0.068	0.021
	[0.071]	[0.071]	[0.080]	[0.081]	[0.080]	[0.081]	[0.067]	[0.066]	[0.059]
Game in São Paulo	0.129	0.129	0.117	0.160**	0.155**	0.101	0.137**	0.126*	0.055
	[0.080]	[0.081]	[0.076]	[0.075]	[0.076]	[0.075]	[0.068]	[0.068]	[0.061]
1 Hour(s) after game		-0.025	0.013		-0.025	0.013		-0.025	0.013
		[0.033]	[0.039]		[0.033]	[0.039]		[0.033]	[0.039]
2 Hour(s) after game		-0.101**	-0.114**		-0.101**	-0.114**		-0.101**	-0.113**
		[0.044]	[0.047]		[0.044]	[0.047]		[0.044]	[0.047]
3 Hour(s) after game			-0.178***			-0.178***			-0.178***
			[0.051]			[0.051]			[0.051]
4 Hour(s) after game			-0.120**			-0.119**			-0.119**
			[0.048]			[0.048]			[0.049]
5 Hour(s) after game			-0.013			-0.013			-0.013
			[0.044]			[0.044]			[0.044]
1 Hour(s) before game		-0.055*	-0.023		-0.055*	-0.023		-0.055*	-0.023
		[0.033]	[0.040]		[0.033]	[0.040]		[0.033]	[0.040]
2 Hour(s) before game		-0.058	-0.064		-0.058	-0.064		-0.058	-0.064
		[0.040]	[0.043]		[0.040]	[0.043]		[0.040]	[0.043]
3 Hour(s) before game			0.02			0.02			0.021
			[0.056]			[0.056]			[0.056]
4 Hour(s) before game			-0.062			-0.061			-0.061
			[0.044]			[0.044]			[0.044]
5 Hour(s) before game			-0.044			-0.044			-0.043
			[0.041]			[0.041]			[0.041]
lnalpha	-3.163***	-3.163***	-3.164***	-3.163***	-3.163***	-3.164***	-3.164***	-3.164***	-3.164***
	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Status interactions	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Other controls (Z_t)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	78,840	78,840	78,840	78,840	78,840	78,840	78,840	78,840	78,840

*Notes: An observation is an hour-long period for the city of São Paulo as a whole over the period 2006-2014. The estimates come from negative binomial regressions; robust standard errors are in brackets. "Lnalpha" is the estimate of the log of the dispersion parameter. Time fixed effects include dummies for the year, month, day of the month, day of the week and hour. Status interactions indicate whether more than one World Cup games are played during the hour-long time block. Other controls (Z_t) include holiday indicators and variables that reflect changes occurred in the crime reporting and policing methods. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE II.A-7: THE IMPACT OF WORLD CUP GAMES ON CRIME IN SÃO PAULO – BY GAMES' OUTCOME

DV: Total crime	(1)	(2)	(3)
Non-Brazil game	-0.055**	-0.067**	-0.119
	[0.026]	[0.034]	[0.076]
Brazil defeat	-0.044	-0.052	-0.061
	[0.083]	[0.083]	[0.083]
Brazil win	-0.048	-0.055	-0.097
	[0.079]	[0.082]	[0.111]
Game in São Paulo	0.143*	0.148*	0.160**
	[0.079]	[0.079]	[0.076]
1 Hour(s) after Non-Brazil game		-0.051	-0.011
		[0.035]	[0.041]
2 Hour(s) after Non-Brazil game		-0.157***	-0.180***
		[0.041]	[0.045]
3 Hour(s) after Non-Brazil game			-0.208***
			[0.057]
4 Hour(s) after Non-Brazil game			-0.159***
			[0.052]
5 Hour(s) after Non-Brazil game			-0.034
			[0.051]
1 Hour(s) after Brazil defeat		-0.099***	-0.108***
		[0.033]	[0.033]
2 Hour(s) after Brazil defeat		-0.102	-0.111
		[0.140]	[0.140]
3 Hour(s) after Brazil defeat			-0.230***
			[0.016]
4 Hour(s) after Brazil defeat			-0.259***
			[0.069]
5 Hour(s) after Brazil defeat			0.034
			[0.022]
1 Hour(s) after Brazil win		0.147**	0.167**
		[0.058]	[0.071]
2 Hour(s) after Brazil win		0.263***	0.254***
		[0.079]	[0.078]
3 Hour(s) after Brazil win			0.031
			[0.134]
4 Hour(s) after Brazil win			0.162**
			[0.081]
5 Hour(s) after Brazil win			0.162***
			[0.044]
1 Hour(s) before Non-Brazil game		-0.036	0.012
		[0.035]	[0.044]
2 Hour(s) before Non-Brazil game		-0.03	-0.025
		[0.044]	[0.048]
3 Hour(s) before Non-Brazil game			0.035
			[0.061]
4 Hour(s) before Non-Brazil game			-0.065
			[0.047]
5 Hour(s) before Non-Brazil game			-0.079*
			[0.046]
1 Hour(s) before Brazil defeat		-0.263**	-0.271**
		[0.133]	[0.133]
2 Hour(s) before Brazil defeat		-0.237***	-0.246***
		[0.044]	[0.044]
3 Hour(s) before Brazil defeat			-0.242***
			[0.044]

Continued	(1)	(2)	(3)
4 Hour(s) before Brazil defeat			-0.205
			[0.155]
5 Hour(s) before Brazil defeat			0.023
			[0.050]
1 Hour(s) before Brazil win		-0.033	-0.013
		[0.076]	[0.073]
2 Hour(s) before Brazil win		-0.1	-0.13
		[0.098]	[0.100]
3 Hour(s) before Brazil win			0.14
			[0.189]
4 Hour(s) before Brazil win			-0.047
			[0.177]
5 Hour(s) before Brazil win			0.038
			[0.092]
lnalpha	-3.163***	-3.164***	-3.165***
	[0.012]	[0.012]	[0.012]
Time fixed effects	Yes	Yes	Yes
Status interactions	No	Yes	Yes
Other controls (Z_t)	Yes	Yes	Yes
N	78,840	78,840	78,840

*Notes: An observation is an hour-long period for the city of São Paulo as a whole over the period 2006-2014. The estimates come from negative binomial regressions; robust standard errors are in brackets. "Brazil win" includes Brazil's wins and draw. "Lnalpha" is the estimate of the log of the dispersion parameter. Time fixed effects include dummies for the year, month, day of the month, day of the week and hour. Status interactions indicate whether more than one World Cup games are played during the hour-long time block. Further controls (Zt) include holiday indicators and variables that reflect changes occurred in the crime reporting and policing methods. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE II.A-8: THE IMPACT OF WORLD CUP GAMES ON CRIME IN SÃO PAULO – PREVIOUS TOURNAMENTS

DV: Total crime	
Previous WCs game (Brazil)	-0.116***
	[0.045]
Game (2014 WC)	-0.055**
	[0.026]
Brazil game (2014 WC)	0.008
	[0.067]
Game in São Paulo	0.143*
	[0.080]
Previous WCs	-0.005
	[0.008]
lnalpha	-3.163***
	[0.012]
Time fixed effects	Yes
Other controls (Z_t)	Yes
N	78,840

*Notes: An observation is an hour-long period for the city of São Paulo as a whole over the period 2006-2014. The estimates come from negative binomial regressions; robust standard errors are in brackets. "Previous WCs game (Brazil)" is a dummy for Brazil games in previous World Cup editions (2006 and 2010). "Game (2014 WC)" is a dummy for 2014 World Cup games. "Brazil game (2014 WC)" is a dummy for Brazil games in the 2014 World Cup. "Lnalpha" is the estimate of the log of the dispersion parameter. Time fixed effects include dummies for the year, month, day of the month, day of the week and hour. Other controls (Z_t) include holiday indicators and variables that reflect changes occurred in the crime reporting and policing methods. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE II.A-9: THE IMPACT OF WORLD CUP GAMES ON CRIME IN SÃO PAULO – ALTERNATIVE TIME FIXED EFFECTS AND SPECIFIC CRIMES

Analysis	Baseline	Kick-off time	Audience	Round	Outcome	Baseline	Kick-off time	Audience	Round	Outcome
DV:			**A) Total crime**					**B) Robbery**		
Game	-0.049*					-0.095***				
	[0.027]					[0.036]				
Game starts at 13hs		-0.01					-0.101			
		[0.041]					[0.064]			
Game starts at 16hs or 17hs		-0.047					-0.085*			
		[0.038]					[0.050]			
Game starts at 19hs or 22hs		-0.11***					-0.102*			
		[0.044]					[0.059]			
WW avg live audience >250M			-0.21***					-0.35***		
			[0.063]					[0.079]		
WW avg live audience <250M			-0.028					-0.067*		
			[0.027]					[0.037]		
Group-stage game				-0.032					-0.065*	
				[0.028]					[0.039]	
Round-of-16 game				0.041					-0.104	
				[0.051]					[0.072]	
Round-of-8 game				-0.36***					-0.33**	
				[0.105]					[0.151]	
Round-of-4 game				-0.25***					-0.25**	
				[0.085]					[0.099]	
Non-Brazil game					-0.050*					-0.1***
					[0.027]					[0.036]
Brazil defeat					-0.008					-0.1***
					[0.084]					[0.063]
Brazil win					-0.041					-0.138
					[0.080]					[0.120]
Brazil game	0.018	0.011	0.151*	0.082		-0.059	-0.066	0.141	0.001	
	[0.069]	[0.072]	[0.081]	[0.068]		[0.097]	[0.101]	[0.104]	[0.103]	
Game in São Paulo	0.131*	0.114	0.150**	0.122*	0.133*	-0.046	-0.049	-0.014	-0.041	-0.048
	[0.079]	[0.080]	[0.073]	[0.067]	[0.079]	[0.108]	[0.112]	[0.103]	[0.101]	[0.107]
lnalpha	-3.2***	-3.2***	-3.2***	-3.2***	-3.2***	-2.8***	-2.8***	-2.8***	-2.8***	-2.8***
	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
Time fixed effects	Altern.	Altern.	Altern.	Altern.	Altern.	Baseline	Baseline	Baseline	Baseline	Baseline
Other controls (Z_t)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	78,840	78,840	78,840	78,840	78,840	78,840	78,840	78,840	78,840	78,840
DV:			**C) Theft**					**D) Assault**		
Game time	-0.047					0.035				
	[0.033]					[0.071]				
Game starts at 13hs		0.048					-0.009			
		[0.044]					[0.095]			
Game starts at 16hs or 17hs		-0.074					0.001			
		[0.046]					[0.114]			
Game starts at 19hs or 22hs		-0.19***					0.125			
		[0.070]					[0.133]			
WW avg live audience >250M			-0.12					-0.165		
			[0.081]					[0.177]		
WW avg live audience <250M			-0.036					0.058		
			[0.034]					[0.074]		
Group-stage game				-0.022					0.035	
				[0.036]					[0.075]	
Round-of-16 game				0.019					0.209	
				[0.056]					[0.168]	
Round-of-8 game				-0.41***					-0.099	
				[0.113]					[0.189]	

Continued	Baseline	Kick-off time	Audience	Round	Outcome	Baseline	Kick-off time	Audience	Round	Outcome
	C) Theft					D) Assault				
Round-of-4 game				-0.161					-0.5***	
				[0.130]					[0.163]	
Non-Brazil game					-0.047					0.037
					[0.033]					[0.071]
Brazil defeat					0.1					-0.036
					[0.075]					[0.236]
Brazil win					-0.031					0.282*
					[0.096]					[0.160]
Brazil game	0.052	0.056	0.113	0.098		0.166	0.196	0.33	0.263*	
	[0.078]	[0.077]	[0.097]	[0.075]		[0.150]	[0.160]	[0.207]	[0.150]	
Game in São Paulo	0.35***	0.32***	0.36***	0.33***	0.35***	-0.085	-0.054	-0.046	-0.029	-0.105
	[0.079]	[0.078]	[0.077]	[0.073]	[0.080]	[0.206]	[0.209]	[0.198]	[0.177]	[0.204]
lnalpha	-3.3***	-3.3***	-3.3***	-3.3***	-3.3***	-3.0***	-3.0***	-3.0***	-3.0***	-3.0***
	[0.020]	[0.020]	[0.020]	[0.020]	[0.020]	[0.031]	[0.031]	[0.031]	[0.031]	[0.031]
Time fixed effects	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Other controls (Z_t)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	78,840	78,840	78,840	78,840	78,840	78,840	78,840	78,840	78,840	78,840

*Notes: An observation is an hour-long period for the city of São Paulo as a whole over the period 2006-2014. The estimates come from negative binomial regressions; robust standard errors are in brackets. "Brazil win" includes Brazil's wins and draw. "Lnalpha" is the estimate of the log of the dispersion parameter. Baseline time fixed effects include dummies for the year, month, day of the month, day of the week and hour. Alternative ("Altern.") time fixed effects include dummies for the year, day of the year, day of the week and hour. Other controls (Z_t) include holiday indicators and variables that reflect changes occurred in the crime reporting and policing methods. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE II.A-10: THE IMPACT OF CONCENTRATION

Panel	Hourly	12-hourly	12-hourly	12-hourly	12-hourly
DV:	Total crime (A)	Total crime (B)	Total crime (C)	Robbery (D)	Theft (E)
Game*Stadium or Viewing PDs	0.595***	0.618***	0.618***	0.221***	0.929***
	[0.144]	[0.046]	[0.046]	[0.082]	[0.060]
Game	-0.066***	-0.151***	-0.147***	-0.192***	-0.164***
	[0.024]	[0.016]	[0.017]	[0.023]	[0.027]
Brazil game	0.005	0.100***	0.114***	-0.031	0.028
	[0.084]	[0.024]	[0.025]	[0.036]	[0.040]
Game in São Paulo	0.123**	0.108***	0.101***	0.041	0.220***
	[0.061]	[0.029]	[0.030]	[0.043]	[0.046]
Time fixed effects	Baseline	Baseline	Alternative	Baseline	Baseline
Other controls (Z_t)	Yes	Yes	Yes	Yes	Yes
N	7,332,120	611,010	611,010	611,010	611,010

*Notes: In column (A) an observation is an hour-long period for one of the 93 PDs in which São Paulo is divided over the period 2006-2014. The estimates come from a Poisson regression with fixed effects; cluster-robust standard errors are in brackets. In columns (B) to (E) an observation is a 12-hour-long period (1 a.m. to midday or 1 p.m. to midnight) for one of the 93 PDs in which São Paulo is divided over the period 2006-2014. The estimates come from negative binomial regressions with fixed effects; standard errors from the observed information matrix are in brackets. Baseline time fixed effects include dummies for the year, month, day of the month, day of the week and hour (Column (A)) or 12-hour period (Column (B), (D) and (E)). Alternative time fixed effects include dummies for the year, day of the year, day of the week and 12-hour period. Other controls include holiday indicators and variables that reflect changes occurred in the crime reporting and policing methods. ***, **, * denote statistical significance at 1%, 5% and 10% respectively.*

TABLE II.A-11: THE IMPACT OF CONCENTRATION: CONLEY AND TABER *VS* STANDARD CONFIDENCE INTERVALS

DV	$\widehat{\delta_1}$	**95% Confidence Intervals**				**N**
		Conley-Taber		Cluster-robust		
		Lower	Upper	Lower	Upper	
Total crime	0.595*º	0.405	0.816	0.313	0.876	7,332,120

*Notes: The coefficient and cluster-robust confidence intervals come from the regression in Table II.A-10, column (A). Conley and Taber's confidence intervals are derived applying the authors' inference method to the same model. * = significantly different from zero at the 5% level using Conley and Taber's approach; º = significantly different from zero at the 5% level using cluster-robust standard errors.*

III.

The effect of police on crime: evidence from the 2014 World Cup in São Paulo

Abstract

This paper estimates the causal impact of police on crime. To address the problem of reverse causality, I consider as a natural experiment the creation of a special police unit to intensify surveillance around a few tournament-related locations in the city of São Paulo during the 2014 FIFA World Cup. The tournament may impact crime in other ways than just through increased policing, namely via fan concentration and voluntary incapacitation. In order to disentangle the specific effect of police on crime, I exploit the fact that the World Cup affected different areas of the city through different channels and at different times. Difference-in-differences estimates reveal that increased police presence leads to a significant reduction in criminal activity.

Keywords: Police • Crime • World Cup • Natural experiment.
JEL classification: K42.

III.1 Introduction

A crucial challenge in the economics of crime literature is to empirically assess the effect of policing on crime. The difficulty arises from the fact that crime and police presence are simultaneously determined, causing a problem of reverse causality. Some authors have tried to break this endogeneity by employing instrumental variables or, more often, natural experiments. A problematic aspect, though, is that most studies fail to acknowledge that the event which triggers the policing change may simultaneously affect criminal activity in other ways. As a consequence, these contributions are likely to measure the overall impact of the triggering event on crime – rather than the specific effect of increased police presence.

This paper tackles the issue by considering an innovative natural experiment and framing it within a conceptual setting that is intended to account for all the effects that the triggering event may have on crime, and ultimately isolate the specific impact of interest. I exploit the natural experiment represented by the creation of a special police unit to monitor specific tournament-related targets in the city of São Paulo during the 2014 Fédération Internationale de Football Association (FIFA) World Cup. The unit was active over a two-month period and consisted of 4,265 extra officers, inflating the police presence in the city by about 25%. The crucial assumption in my empirical strategy is that the selective (across neighborhoods) police increase is exogenous with respect to the underlying evolution of crime, thus representing a natural experiment that breaks the simultaneous determination of crime and policing.

I use data on reported crimes that were committed in the city of São Paulo throughout the period 2006-2014, aggregated daily and by police district. I focus on the total number of crimes, robberies and thefts. I link these data to information on the World Cup tournament and the deployment of the special police unit.

A difference-in-differences analysis to compare criminal outcomes in districts with increased protection with respect to other districts would likely provide an estimate of the overall effect of the World Cup on crime. I adopt a more refined approach by employing a conceptual framework to understand the World Cup-crime relationship. The tournament can be expected to affect local criminal behavior in three ways: (a) increased police presence reduces crime; (b) higher concentration positively affects the incidence of occurrences; and (c) the voluntary incapacitation of a substantial number of individuals who are watching the games reduces

criminal interaction. In my research strategy, I exploit the fact that the World Cup affects different areas of the city in different ways and at different times to root out the voluntary incapacitation and concentration effects. Then, the difference-in-differences estimates capture the specific impact of police on crime.

My results show that the police have a negative and significant impact on crime: the total number of offenses per day in the city of São Paulo decreases by 18%, while the number of robberies drops by 34%. Results show that this effect is rather local as neighboring districts do not experience crime reductions. In addition, I find no evidence that offenders from the treated districts are displacing their illegal activities to surrounding areas, with the exception of thieves, who seem to be spatially diverting offenses to some extent. Additional tests suggest that there is no significant temporal displacement, so that the observed crime reductions actually represent prevented – rather than temporally displaced – offenses.

To enhance confidence that I am actually capturing the causal effect of police on crime, I calculate, as placebo treatments, the difference-in-differences estimates for the period before the policy-on one in 2014 and for the periods corresponding to the policy-on one in previous years. Results validate my exercise in that they reveal no special crime dynamics affecting the treated districts in these pseudo-treatment periods.

The literature acknowledges two channels through which police presence may reduce crime: deterrence and incapacitation. On the one hand, the presence of the police makes criminal activity more costly (deterrence). On the other hand, more officers on the streets likely results into more offenders behind the bars (incapacitation). I believe my results are best interpreted as exclusively capturing the deterrence effect. In fact, the natural experiment considered in this paper (based on more police on the street), involves the deployment of a clear deterrence strategy. Moreover, the policy-on period was short (2 months), and it is unlikely that increased incarceration could be made effective and start producing a significant impact so quickly.

The remainder of this paper is organized as follows. Section I.2 reviews the literature on the effect of police in crime. Section III.3 describes the World Cup-related police increase in São Paulo. Section III.4 provides a conceptual framework to understand the World Cup-crime relationship. I introduce the data and describe my empirical strategy in Section III.5. Section III.6 presents my baseline results as well as a number of sensitivity and robustness checks. Section III.7 concludes.

III.2 Literature review and contribution

III.2.1 Previous literature on the effect of police on crime

This contribution is linked and adds to previous research aimed at assessing the effect of policing on crime. The most influential theoretical result on the expected effect of increased police presence on crime can be found in Becker (1968). In his seminal contribution, Becker follows "the economists' usual analysis of choice" and takes criminals as rational beings who weight the benefits and costs of their actions. A core conclusion is that a rise in the probability of conviction, achieved for instance through an increase in police presence, reduces crime by lowering the expected utility from committing an offense – that is, through a deterrence effect. Subsequent studies (for instance, Ehrlich, 1981) have stressed that increased police activity can also reduce crime by taking offenders out of circulation via detention or incarceration – that is, through an incapacitation effect. In sum, theory suggests that there are two channels through which increased police presence may affect crime: deterrence and incapacitation. Either way, the expected impact is negative.

Although this outcome is rather straightforward in theory, empirically estimating the causal effect of police on crime has proven to be challenging because of a major endogeneity issue. While police presence is expected to (negatively) affect crime, crime is likely to (positively) impact police presence, as more dangerous neighborhoods are usually allocated more officers. In order to identify the effect of police on criminal activity it is necessary to break this endogeneity. Some authors have tried to do so by employing instrumental variables or, more often, natural experiments.

Levitt (1997) provides evidence that expansions in the police force in the US are disproportionately concentrated in mayoral and gubernatorial election years. Considering a panel of 59 large US cities over the period 1970 to 1992, he uses the timing of elections to instrument for changes in the size of the police force. His findings point to a negative and significant effect of police on violent crime. An issue with Levitt's approach is that the timing of the elections may impact crime in other ways than through police presence, for instance by affecting the sentencing behavior of judges. In addition, McCrary (2002) shows that, after adjusting for a computational error, Levitt's results are no longer statistically significant (see also Levitt's reply, 2002).

Di Tella and Schargrodsky (2004) gave start to a line of studies that use terrorism-related events to look at the police-crime relationship. The rationale is that (the fear of) terrorist attacks on specific targets induce(s) exogenous variations in the allocation of police resources, thus breaking the simultaneous determination of crime and police presence. Under certain conditions, this allows for the estimation of the causal impact of policing on crime. Di Tella and Schargrodsky (2004) consider a terrorist attack to an important Jewish center in Buenos Aires occurred in July 1994. Following this event, the Argentinean government assigned police protection to every Muslim and Jewish institution in the country. The authors look at data on the number of motor vehicle thefts per block in three Buenos Aires neighborhoods over the period April 1994 to December 1994. Using a difference-in-differences estimator, they find that auto theft declined by 75% on blocks that received extra protection, compared to other blocks. This corresponds to an approximated elasticity of crime to police presence of -0.33.

Klick and Tabarrok (2005) exploit the exogenous shocks in police presence in Washington, D.C., pursuant to changes in the terror alert level set by the US Department of Homeland Security. The authors conclude that an increase in police presence of about 50% leads to a significant crime reduction on the order of 15%.

Draca, Machin and Witt (2011) consider London's terror attacks of July 2005. Following the occurrences, police presence was substantially increased in some of the city's central boroughs for a limited period of time. Exploiting detailed data on crime and police deployment, the authors estimate an elasticity of crime with respect to police of approximately -0.3 to -0.4.

A potential problem with studies that use terrorism-related events to look at the crime-police relationship is that of correlated shocks. Terror likely affects crime in more ways than just through increased policing, and there is no clear framework to account for such complex relationship. In particular, a heightened threat level could make potential victims and criminals fear for their own safety, and thus avoid areas or times in which there is an increased perceived risk of being harmed. Then, the potential for criminal interactions would be reduced. Absent a well-defined framework, this (and any other) effect cannot be disentangled from that of increased policing. Then, the estimated coefficients may ultimately be biased, as they reflect the overall impact on crime of higher terror levels rather than the specific effect of increased policing.

Ideally, one would like to consider a policing change whose triggering event is unlikely to affect crime through channels other than the change in policing. This is the case of Poutvaara and

Priks (2009). The authors analyze the impact of policing on hooligan violence by considering the sudden reallocation of the Stockholm special police unit that targets hooliganisms to different activities following the 9/11 terrorist attack in September 2001 and the Tsunami catastrophe in December 2004. The key point here is that local hooligan violence is unlikely to be affected by the 2004 Tsunami or the 9/11 attacks in other ways than through the policing change. Difference-in-differences analysis reveals that Stockholm-related hooligan violence increased dramatically during these periods, implying that policing has a negative and substantial effect on this kind of crimes.

Natural experiments like the one exploited by Poutvaara and Priks (2009) are rare to come across. An alternative way to deal with the correlated shocks issue is – when possible – to set the analysis within a conceptual framework that makes it possible to understand and account for the complex relationship between the triggering event and crime, with the ultimate purpose of isolating the specific impact of the policing change. This paper follows such approach.

III.2.2 Contribution to the literature

Although closely related to the discussed literature, this study provides an original contribution to the understanding of the effect of increased police presence on crime. To begin with, I propose a novel approach for dealing with the correlated shocks issue. A well-defined conceptual framework for the World Cup-crime relationship allows me to isolate the impact of interest by rooting out the other channels through which the event may impact delinquency.

Furthermore, for a change, I do not consider a terror-related increase in policing. This makes the comparison between my results and those from previous studies more interesting. The effect on crime of an exogenous police increase should not depend on the kind of event that triggered it. Thus, while moderate dissimilarities in estimated effects can be explained by reason of the different settings, large differences may raise some concerns.

Moreover, to the best of my knowledge this is the first study that looks at the Brazilian context. Ranking tenth in the World Bank list of countries by intentional homicide rate, Brazil can be considered as one of the most dangerous not-at-war states in the world. In addition, the country has an alarming record of police violence (see *The Economist*, 2014), typically triggering high levels of "revenge" violence (see *Folha de S. Paulo*, 2015), in a vicious circle that has proven hard to break. In this context, most people seem to have lost trust in the force that should

be protecting them (see *BBC*, 2014). These (sad) peculiarities make Brazil an especially interesting ground to study whether and to which extent increasing police presence is an effective tool in the fight against crime, and whether this remains true in a high-violence, weak-institutions context.

Finally, I use daily data which are less subject to endogeneity problems from crime to police and allow for a more in-depth control of seasonality issues.

III.3 The World Cup-related police increase in São Paulo

The 2014 edition of the FIFA World Cup – the quadrennial world championship for men's national football teams organized by FIFA – took place in Brazil from June 12th 2014 to July 13th 2014, and ended with Germany’s victory. Thirty-two national teams played a total of 64 games in 12 venues located in as many host cities across Brazil, including São Paulo.

The State of São Paulo created a special police unit, the *Comando de Policiamento Copa* (“CPCopa”), to provide additional surveillance in the city on the occasion of the World Cup. The unit counted with three battalions for a total of 4,265 policemen, inflating the police presence in the city by about 25%. The CPCopa was not distributed evenly around the city. On the contrary, it was exclusively assigned to specific spots located in 11 out of São Paulo’s 93 police districts (“PDs”), as shown in Table III.A-1. I refer to these 11 districts as “target PDs”. The monitored spots comprised:

- Venues of World Cup-related events: these include the locations of official celebrations and maxi-screens for public views of the games. The most popular of these venues was by far the FIFA Fanfest, which could accommodate up to 40 thousands viewers. The FIFA Fanfest was located in the city center (*Vale do Anhangabaú*).
- Public transport and touristic spots: these include the main touristic attractions in the city (such as museums, theaters and high streets) and some crucial transportation spots.
- FIFA hotels: these include the eight hotels where the teams and FIFA representatives were hosted.

- Training centers: in view of the tournament, FIFA had pre-selected four training centers as suitable within the city of São Paulo. However, only one of them was actually chosen as a base camp by a national team: the US squad used the São Paulo training center as its preparation site. The center is located in the northern part of the city (*Barra Funda*).
- The stadium: The São Paulo Arena was built between 2011 and 2014 to host the World Cup matches in São Paulo. It is located in a peripheral region, in the north-eastern part of the city (Zona *Leste*) and nearly 19 km away from the center. It can accommodate about 63,000 viewers. The CPCopa was responsible for ensuring safety outside the stadium and in the surrounding area – while a private service was employed by FIFA to grant security inside the Arena.

In the deployment of the CPCopa, caution was paid so that the regular levels of police presence would be maintained unchanged in the rest of the city. To this end, the government hired more than 2,000 new recruits and managed the holidays of the existing officers, so that the CPCopa was actually made up of *extra* (rather than re-allocated) force. The unit was active from May 20th 2014 (24 days before the World Cup opening) until July 20th 2014 (one week after the end of the tournament) – but not all targets were monitored throughout the whole period. I refer to this time window as "the CPCopa period". Following the dissolution of the CPCopa, the officers were re-allocated across the whole State of São Paulo.

For the purposes of this analysis, it is important to describe some characteristics of the surveillance in the areas of the stadium and of the São Paulo training center ("TC"). The special police unit supervised the proximities of the São Paulo Arena 24 hours per day on each day during the CPCopa period. On days in which a game was played at this stadium, 610 CPCopa policemen divided into three daily shifts watched over the area. On other days, 90 policemen (in three daily shifts) were employed. The TC did not receive increased policing throughout the CPCopa period, but only over a shorter period. The US team arrived in São Paulo on June 9th 2014 and was eliminated by Belgium in a round-of-16 game on July 1st 2014. An average of 25 CPCopa policemen (divided into two daily shifts) watched over the TC 24 hours per day during the period June 8th to July 2nd 2014. Figure III.A-1 describes the timeline of the CPCopa activity, with special regard to the surveillance at the São Paulo Arena and TC.

III.4 Conceptual framework: the World Cup-crime relationship

The World Cup is likely to affect crime in other ways than just through increased police presence. These mechanisms need to be considered and somehow netted out in order to isolate the specific impact of police on crime. Marie (2016) considers information on football matches for nine London teams linked to crime data at the borough level. He introduces a framework to analyze the multiple effects of sporting events on crime, and develops some simple assumptions to disentangle and estimate them. According to Marie’s framework, large sporting events can be expected to affect local crime in three ways: (a) the concentration of hostile fans increases the incidence of violent offences; (b) the displacement of police personnel sent to monitor the event positively impacts crime in the areas from which the police were displaced; and (c) the voluntary incapacitation of a substantial number of individuals who are attending the event reduces criminal activity.

Marie’s framework is a valuable starting point for understanding the World Cup-crime relationship. However, the framework for the present study differs in some aspects with respect to Marie’s one. First of all, there was no displacement of police officers in the city of São Paulo during the World Cup: a special unit made up of *extra* force was created. Therefore, the police displacement effect does not apply to this study. In line with the literature review presented in the previous section, the expected impact of increased police presence on crime is negative. Concentration and voluntary incapacitation are relevant to the World Cup-crime setting, though with some dissimilarities with respect to Marie’s framework. I discuss each of these effects and their likely impact below.

Concentration

It is rather straightforward that a huge event such as the World Cup leads to concentration in the cities hosting the games. It was estimated that during the month-long event Brazil received about 1 million foreign visitors from 203 countries (the corresponding figure for June 2012 was 0.3 million), as well as about 3 million domestic tourists. The city of São Paulo alone received about 540 thousands tourists, 220 thousands of whom were foreigners.

It is reasonable to expect that concentration affects criminal activity. In fact, concentration (or population density) is ordinarily considered as a crime-affecting factor in the literature (for instance, Kelly, 2000). Higher concentration increases the likelihood of interactions between potential victims and offenders, and may negatively impact the chance of being caught (Kelly, 2000). As a consequence, everything else being equal, concentration is expected to have a positive effect on criminal activity.

This outcome is supported by empirical results. For instance, Campaniello (2013) uses a yearly panel of crime data for Italian provinces to study the overall effect of increased tourism on crime. She exploits the 1990 FIFA World Cup as a source of exogenous variation in the level of attractiveness of different provinces and finds that hosting such an event leads to a significant increase in property crimes.

Marie (2016) argues that increased concentration of hostile fans should result in more violent offenses. However, in the context of the World Cup, it is likely that hostile fans represent a smaller fraction of the public than in local championships. In addition, Campaniello's (2013) results reveal a significant effect of World Cup-related concentration on property crimes. As a consequence, differently form Marie (2016), I expect concentration to impact both property and violent crime.

Voluntary incapacitation

The World Cup games are one of the most followed sporting events in the world. Viewership is made up not only of stadium-goers (54 thousands per match on average in 2014) but also (mostly) of people who watch the match on TV or online. FIFA estimates that the 2014 competition reached 3.2 billion in-home viewers worldwide – little short of half the world population – and 1 billion people watched the final. It is hard to imagine an event that could get a comparable audience record. The voluntary incapacitation effect has to do with the fact that, while so many people are busy watching the games, the potential for criminal interactions drops.

The term "incapacitation" has traditionally been used in the crime literature to express that those who are incarcerated or otherwise restrained in their freedom to act cannot commit offenses (for instance, Ehrlich, 1981). More recently, some studies have introduced the concept of "voluntary incapacitation", pursuant to which potential offenders and victims voluntarily engage in some activities, thus substituting away from alternative occupations that would more likely

lead to crime. Then, the expected effect of World Cup games on crime through voluntary incapacitation is negative.

The literature mostly emphasizes the role of voluntary incapacitation on potential criminals. However, it is important to remark that there can also be a parallel effect on potential victims, who may become more or less susceptible to attacks while they engage in a certain activity with respect to their alternative occupation. Although it is often hard to distinguish between the impacts on the two sides of criminal interaction, the interpretations of the results and the policy implications remain essentially unchanged.

It has been shown that movie attendance (Dahl and Della Vigna, 2009) and video game playing (Cunningham, Engelstätter and Ward, 2016) generate a significant voluntary incapacitation effect: violent crime in the short term is reduced as potential criminals (and victims) choose to go to the movie theater or play video games instead. More closely related to the World Cup context, Laqueur and Copus (2014) analyze local criminal activity in Chicago during major sporting events. Their results indicate substantial reductions in crime during the hours in which important sporting events are broadcast on television. Finally, analyzing football matches in London, Marie (2016) finds that property crime decreases for away games and interprets this as a result of voluntary incapacitation of potential offenders attending the matches.

To sum up, the World Cup is expected to affect local crime rates through three channels: police increase, voluntary incapacitation and concentration. Table III.A-2 portrays the expected direction of the impact of each mechanism.

III.5 Data and empirical strategy

III.5.1 Data

I exploit the natural experiment represented by the creation of a special police unit aimed at monitoring strategic targets in São Paulo during the 2014 FIFA World Cup to investigate the effect of increased police presence on crime. I use a panel of daily data on reported crimes across the 93 police districts in which the city of São Paulo is divided, and combine them with information on the CPCopa deployment and World Cup matches.

I obtained crime data from the Public Safety Department of the State of São Paulo (*Secretaria de Estado da Segurança Pública do Governo de São Paulo*). These entail crimes reported to the police that occurred in the city of São Paulo in the period 2006-2014. The data portray the criminal incidents in which the police completed a case report. They cover common offenses which I aggregated into six broader crime type categories, namely: murder (committed and attempted), robbery, rape, assault, theft and drugs-related offenses. All records provide information on the date of the occurrence as well as on the police district (*distrito policial*) where it took place.[6] There are 93 standard PDs in São Paulo. I organized crime records in a panel by day and PD.

Error! Reference source not found. contains some descriptive statistics about the data I am using. It can be noticed that, as it is often the case with high-frequency criminal occurrences, the distributions are positively skewed.

III.5.2 Research design

My purpose is to estimate the causal effect of increased police presence on crime. I explore a panel of daily data on criminal occurrences in each of the 93 PDs in the city of São Paulo over the period January 1st 2006 to December 31st 2014. The increase in police presence I analyze is represented by the creation of a special unit, the CPCopa, aimed at providing additional policing in São Paulo on the occasion of the 2014 FIFA World Cup. The CPCopa was active over the period May 20th 2014 to July 20th 2014 (although not all targets were monitored throughout this whole period), so it operated from 24 days before to one week after the World Cup. The special unit was exclusively assigned to the monitoring of specific World Cup-related spots, which were located in 11 different PDs (the target PDs).

For the sake of clarity, my conceptual starting point is a definition of the treatment group as including all target PDs and the treatment period as comprising, for each target PD, the whole period over which the CPCopa was active in monitoring it. This setting is naïve because it does not take into account that crime outcomes during the World Cup are also affected by increased concentration and voluntary incapacitation. Thus, this approach could at best provide an estimate of the overall effect of the World Cup on crime – rather than the specific impact of increased

[6] When requesting data from the SSP, I asked for the exact location of the occurrences. However, it was not possible to obtain this level of detail.

policing. I refine the treatment group and period definitions to build a design that makes it possible to isolate the effect of interest. In particular, I root out the voluntary incapacitation and concentration effects in turn by exploiting the fact that the World Cup affects different areas of the city in different ways and at different times. For this purpose, I discuss the temporal and geographical reach of the voluntary incapacitation and concentration effects.

The voluntary incapacitation effect implies that during World Cup matches the number of offenses decreases as potential offenders and victims are engaged in watching the games, so that the probability of criminal interaction is lower. In case this effect was not filtered out, the negative impact of police on crime would be overestimated.

As regards the geographical dimension, I expect voluntary incapacitation to be at play throughout the city, as most people watch the game on TV (or online). On the other hand, concerning the temporal dimension, this effect is only present during matches. This allows me to easily root this mechanism out by excluding from the sample those 25 days on which 2014 World Cup matches were played. Given that the CPCopa was active for a longer period than the tournament itself, this sample cut still leaves me with 37 days over which no tournament game was played and the CPCopa was operational.

The arrival of tourists attracted by the World Cup led to higher concentration in São Paulo. Concentration is expected to positively affect crime by increasing the probability of criminal interaction and reducing the chance for an offender of being caught. In case this factor was not cleared out, the negative impact of police on crime would be underestimated.

The temporal and geographical reach of the concentration effect is not as clear-cut as that of the voluntary incapacitation effect. Tourists may arrive before the official beginning of the tournament and leave after its conclusion. Also, although concentration is most likely to be higher around touristic and World Cup-related spots, it is not obvious how exactly to draw a line. Luckily, the CPCopa setup itself is very telling about the temporal and geographical reach of the concentration effect. In fact, the concern with the potential adverse effects of concentration was at the basis of the creation of this additional police unit. Arguably, the policy-maker studied which areas were likely to be affected the most by the event, and over which period, and arranged the CPCopa's setup accordingly. I rely on the policy-maker's conclusions and expect the concentration effect to be more relevant in target PDs and over the CPCopa period.

Filtering out this effect requires some considerations on the likely timing of the increased concentration across different types of CPCopa target spots. In particular, I consider, for each kind of target spot, its likelihood of attracting higher-than-normal concentration on game days *versus* non-game days, during the CPCopa period.

- Venues of World Cup-related events: Although the events were only scheduled during game days, the FIFA Fanfest was located in a central, touristic area (*Vale do Anhangabaú*). Therefore, it is likely that concentration was high even on non-game days, as tourists took the chance to visit the city.
- Public transport and touristic spots: these were likely to attract higher-than-normal concentration on both game and non-game days.
- FIFA hotels: Both on game and non-game days a crowd may concentrate around hotels to catch a glimpse at the teams or FIFA VIPs leaving or arriving at the premises.
- The stadium: The São Paulo Arena obviously attracted big crowds on days on which a game was played at that venue (six matches took place in São Paulo, including the opening one). However, the location of the stadium makes it unlikely that concentration would be higher-than-normal on other days. The São Paulo Arena is located in a peripheral region and there are no significant touristic attractions in the proximities of the stadium, nor is the region known for its nightlife. In fact, the choice of the stadium's location was mainly driven by the policy-maker's desire to stimulate economic activity in a highly-populated and under-developed area of the city. Bearing these considerations in mind, it is reasonable to expect that concentration around the stadium was not higher-than-normal during non-game days. This implies that, on such days, the concentration effect did not apply to the stadium area.
- The São Paulo TC: The center is located in a region that does not offer touristic attractions and is commonly perceived as dangerous. On both game and non-game days, the TC area is unlikely to attract higher-than-normal concentration.

In sum, the concentration effect is likely to apply to the whole CPCopa period and all CPCopa target spots (and corresponding target PDs). The only exceptions are represented by the stadium and the TC areas, where the concentration effect arguably did not apply during non-game days. Exploiting this fact, I can root out the concentration effect by dropping from the sample all target PDs but the stadium and TC ones. This way, the treatment group is made up of only two

PDs – the stadium and the TC PDs –, while the control group includes all of the 82 non-target PDs.

Conceptually starting from a naïve delineation of the treatment group and period as including, respectively, all target PDs and the whole period over which the CPCopa monitored each PD, I have:

- Rooted out the voluntary incapacitation effect: by restricting the treatment period to only include those days over the CPCopa period on which no World Cup matches were played; and
- Rooted out the concentration effect: by restricting the treatment group to only include those target PDs where no concentration effect is expected on non-game days.

Error! Reference source not found. elaborates in more detail the contents of Table III.A-2 to graphically summarize my research design and highlight the relevant groups for comparison.

III.5.3 Empirical model

My purpose is to estimate the causal effect of increased police presence on crime. I start from a panel of daily data on criminal occurrences in each of the 93 PDs in the city of São Paulo over the period 1st January 2006 to 31st December 2014. According to the research design set out above, I drop from the sample all observations for the 25 days on which 2014 World Cup games were played, as well as observations for all target PDs but the stadium and the São Paulo TC PDs.

The stadium and TC PDs constitute the treatment group, while the 82 non-target PDs represent the control group. I include a rich set of time fixed effects to make sure that all common shocks in the evolution of crime across districts are absorbed. I also include PD indicators to control for unobservable determinants of crime that are invariant at the district level. I obtain the difference-in-differences ("DID") estimator of the effect of police on crime using the following model:

(III.1) $Crime_{s,t} = \beta_0 + \beta_1(CPCopa_t * Stadium_s + CPCopaTC_t * TC_s) + \beta_2 CPCopa_t + \beta_3 CPCopaTC_t + \beta_3 WC_t + \beta_5 ArenaSP_t + Z_t + \rho_t + \theta_s + \varepsilon_{st}$

Where the subscripts s and t denote PD and date, respectively; $CPCopa$ is a dummy equal to one during the CPCopa period (May 20th 2014 to July 20th 2014); $Stadium$ is a dummy equal to one for the stadium PD; $CPCopaTC$ is a dummy equal to one during the period over which the CPCopa monitored the TC (June 8th 2014 to July 2nd 2014); TC is a dummy equal to one for the TC PD; WC is a dummy equal to one during the 2014 World Cup period (June 12th 2014 to July 13th 2014); $ArenaSP$ is a dummy that equals one for the period after the Arena São Paulo's inauguration (which occurred on May 18th 2014); ρ is a set of time fixed effects including dummies for the year, month, day of the month and day of the week; θ are PD fixed effects. Finally, Z comprises a holiday indicator and a dummy variable equal to one starting from December 1st 2013, when it became possible to file reports for robbery online.

I use as baseline dependent variables, $Crime$, the natural logarithm of the total number of crimes, the natural logarithm of the number of robberies and the natural logarithm of the number of thefts. This is because criminal incidents are positively skewed and, for these offenses, the frequency of day/PD combinations with zero reported crimes is low (0.2% of cases for total crime, 2.8% for robbery and 5.3% for theft – see Table III.A-3). On the other hand, any offense other than the three above presents a non-trivial number of zero count observations (see Table III.A-3), making it unfeasible to use the logs.

The identification strategy relies on two key assumptions. The first is that the allocation of the CPCopa to the stadium and TC PDs (the treatment group) is exogenous with respect to the underlying evolution of crime. Officers were placed in those areas in view of the World Cup rather than in response to a change in local criminal activity. In this sense, the CPCopa deployment provides a natural experiment that breaks the simultaneous determination of crime and police presence. The second crucial assumption is that my research design makes it possible to effectively root out the other mechanisms through which the World Cup affects local crime, and isolate the specific impact of increased police presence.

Figure III.A-2 graphically summarizes my DID approach focusing on year 2014. Overall, delinquency has been decreasing across the city in 2014. It can be noticed that the dynamic of criminal activity in the stadium and TC PDs is reasonably close to that in non-target PDs. However, there is some visual evidence that during the CPCopa period the former districts

underwent a crime reduction that appears specific and more extreme with respect to the other districts. This paper argues and motivates that such deviation can be attributed to the increased police presence in the stadium and TC PDs.

A complication in this setup is represented by the fact that I only have two treated PDs. Using N_1 to denote the number of treatment units and N_0 to denote the number of control units, I have N_1=2 and N_0=82. Standard inference methods used in DID models may not perform well in this case, because they are based on the asymptotic approximation that both N_1 and N_0 are large. In particular, Conley and Taber (2011) show that if N_1 and the number of periods (T) are fixed, then the DID estimator is unbiased but inconsistent, so that the estimated treatment coefficient tends in probability to the true parameter of interest plus a noise ($\widehat{\beta_1} \rightarrow \beta_1 + W$). In this case, using the standard inference methods would result in misleading standard errors. The authors develop an alternative approach to inference under the assumption that N_1 is small (finite), using asymptotic approximations that let N_0 grow large. The key idea is to use information on the residuals of the control group to estimate the distribution of the noise W. This approach allows for the calculation of reliable confidence intervals for the treatment coefficient. In my baseline analysis, I use the Conley and Taber's inference method to calculate confidence intervals for the coefficient of interest (β_1).

III.6 The effect of police on crime

III.6.1 Results

Panels (A) and (B) in Table III.A-5 report the results from the estimation of Equation (III.1) using my baseline dependent variables of interest (one for each row). Next to the DID point estimates for β_1, I display the 95% confidence intervals obtained from using the Conley and Taber's inference method. In all regressions, I include fixed effects at the PD level and further controls as displayed in Equation (III.1). In addition, model (A) comprises time fixed effects for the year, month, day of the month and day of the week. Model (B) contains time fixed effects for the year, day of the year and day of the week.[7,8]

---

[7] I drop February 29th in leap years.
[8] For completeness, I present estimation outcomes for total and specific crimes in levels in Table III.B-2.

The results show that increased police presence generates a significant reduction in total crime and robberies, while there is no detectable impact on thefts. Outcomes from models (A) and (B) are virtually identical, in terms of both point estimates and significance considerations. This suggests that results are robust to the choice of the included time fixed effects. In the rest of the paper, I use as baseline time fixed effects those employed in specification (A).

The impact of increased police presence on total crime and robberies is quite large. The DID estimates indicate that the total number of crimes per day went down by 18% and the number of robberies per day dropped by 34% during the treatment period. During such period, in non-target PDs there were on average 12.1 crimes a day, 5.3 of which were robberies. Then, a back-of-the-envelope calculation indicates that the increased police presence prevented on average 2.2 crimes per day, including 1.8 robberies. Over the 62-day period in which the CPCopa was active, about 136 crimes were averted, including 112 robberies.

As mentioned, the literature acknowledges two channels through which police presence reduces crime: deterrence and incapacitation. On the one hand, the presence of the police makes criminal activity more costly (deterrence). On the other hand, more police on the streets likely results into more offenders behind the bars (incapacitation). Like Di Tella and Schargrodsky (2004), Klick and Tabarrok (2005) and Draca, Machin, and Witt (2011), I believe my results are best interpreted as exclusively capturing the police deterrence effect. In fact, the natural experiment considered in this paper (based on more police on the street), involves the deployment of a clear deterrence strategy. Moreover, the CPCopa period was short (2 months), and it is unlikely that increased incarceration could be made effective and start producing a significant impact so quickly.

To calculate the elasticity of crime with respect to police, I need to compute the percentage increase in police presence in the stadium and TC PDs. Data on police presence per PD are not available as they are considered sensitive and confidential. As an approximation, I assume that the 17 thousands policemen active in São Paulo in 2014 were uniformly allocated across the 93 PDs. Then, before the CPCopa, each PD counted with 183 officers. Assuming that they work eight-hour shifts and an average of 21 days per month, there are approximately 43 officers on patrol in each PD at any given time – 86 in the stadium and TC PDs jointly. The CPCopa force increased this number to 128 (30 extra officers in the stadium PD and 12 extra officers in the TC

district at any given time). Thus, the approximate percentage change in police presence in treated districts is 49%, yielding an elasticity of crime with respect to police of -0.37.

This number is remarkably close to previous estimates. Klick and Tabarrok (2005) calculate an elasticity of -0.3 in Washington, D.C., and Draca, Machin, and Witt (2011) estimate it at -0.3 to -0.4 in London. Di Tella and Schargrodsky's (2004) estimate for Buenos Aires is also close (-0.33), but it specifically refers to car thefts. The proximity of elasticity estimates across studies suggests that the deterrent effect is surprisingly robust across contexts: both in London (about one murder per 100,000 people in 2014) and in São Paulo (about 10 murders per 100,000 people in 2014), a 10% increase in police presence leads to a 3-4% reduction in crime. In addition – and quite reasonably – the impact of increased police presence seems not to depend on the nature of the even that triggered it – whether it was a terrorist act or the World Cup, extra officers appear to have the same effect on crime. Nonetheless, caution is in order as regards my calculation of the elasticity, as it is based on the strong assumption that, apart from the CPCopa officers, policemen in São Paulo are uniformly allocated across the 93 PDs.

Panel (C) in Table III.A-5 shows the estimation results from using a definition of the treatment group which comprises all (eleven) target PDs. Notice that in this case N_1=11, so I rely on standard, cluster-robust standard errors. As explained, in all target PDs but the stadium and TC ones, World Cup-related concentration is likely to affect criminal activity even on non-game days. Then, this approach is naïve because it does not allow netting out the impact of concentration. Based on my conceptual framework, I expect this to set back the (negative) impact of increased police presence on crime. Results are consistent with this prediction. Estimated coefficients for total crime and robbery are negative but smaller in absolute value, and only the latter is significant. The coefficient for theft becomes positive, though not precisely estimated.

Finally, I report in panel (A) in Table III.B-1 the results from estimating Equation (III.1) using data from all 93 PDs while including a dummy for the interaction between non-treated target PDs (all target PDs but the stadium and TC ones) and the CPCopa period. In panel (B) in Table III.B-1 I display the results from estimating Equation (III.1) using a sample that comprises all days in the period 2006-2014 and including additional indicators for game days, days when a match was played in São Paulo and a dummy for the interaction between the latter variable and the stadium PD. Results are robust to these alternative specifications.

III.6.2 Robustness and specification checks

In this section, I present additional tests to assess the validity of my results exploiting the spatial and temporal characteristics of the CPCopa operation.

III.6.2.1 Spatial issues: displacement and spill-over

The CPCopa was assigned to the monitoring of specific World Cup-related targets. Thus, one would expect the effect of the police unit to be rather local and well-delimited as regards its spatial reach. My empirical approach, indeed, uses all non-target PDs as control regions, implicitly assuming that they were unaffected by the increased police presence in the stadium and TC PDs during the CPCopa. If this was not the case, my DID estimates would be biased. Thus, it is useful to assess the validity of my assumption.

The increased police presence in the treated PDs may have impacted criminal activities in other areas in two opposite ways. On the one hand, the change in relative police presence across districts may have pushed criminals to displace their illegal activities from treated PDs to other regions, thus causing crime to rise in such regions. On the other hand, the negative effect of police may have spilled over to other PDs, causing crime to fall in such regions. Either way, it is reasonable to assume that the affected areas (if any) would be the PDs adjacent to the treated ones. I examine the two possible effects separately.

If spatial displacement was at play, my DID approach would overestimate the real impact of police on crime. In fact, this effect translates into increased criminal activity in the non-treated PDs, resulting into a larger (in absolute value) DID baseline estimate. Following Draca, Machin, and Witt (2011), I test for spatial displacement by running a robustness check where the control group is restricted to the set of non-target PDs that are adjacent to the stadium and TC districts. This comprises 11 districts. If crime were diverted to these areas, then I would obtain bigger (in absolute value) DID estimates with respect to my baseline results.

As displayed in Table III.A-6, spatial displacement does not seem to be an issue as far as total crime and robberies are concerned. In fact, the police effect estimated using only neighboring PDs as a control group is very close to that in my baseline analysis. On the other hand, there is evidence that some share of thefts is being diverted from the districts receiving extra monitoring to the surrounding ones. The estimated coefficient is significant and about one

third bigger, in absolute value, than the corresponding figure in my baseline results. This suggests that localized police presence displaces theft to relatively less monitored districts.

The negative effect of police may be spilling over to PDs other than the treated ones, causing crime to fall. In this case, my DID approach would underestimate the real impact of police on crime. In fact, this effect translates into lower criminal activity in the non-treated PDs, resulting into a smaller (in absolute value) DID baseline estimate. I analyze this issue by using Equation (III.1) to compare crime outcomes in the 11 districts that surround the stadium and TC PDs (the pseudo-treatment group) against the remaining 71 non-target PDs (the pseudo-control group). Notice that in this case N_1=11, so I rely on the standard, cluster-robust standard errors. If the police effect was spilling over to neighboring PDs, then I would obtain negative and significant DID estimates.

As displayed in Table III.A-6, spill-over effects do not seem to be an issue. Increased police presence in the stadium and TC PDs has no significant impact on total crime and robberies in the neighboring districts. The positive and significant effect on thefts does not reflect a spill-over dynamic (it would be negative in that case), but rather shows that some portion of thefts is being diverted from the monitored PDs to the surrounding ones, in line with results from the spatial displacement analysis.

III.6.2.2 Temporal issues: placebos and displacement

My baseline results indicate that during the CPCopa period criminal activity in the treated PDs decreased significantly more than in non-target ones. Although my research design is aimed at capturing the impact of increased police, I cannot explicitly rule out that the crime reduction is actually driven by a contemporaneous shift in some unobservable factors which I am not properly accounting for. In such case, my DID estimates would be capturing a spurious correlation. I perform some placebo checks exploiting the temporal characteristics of the CPCopa program to alleviate this concern.

To begin with, I check whether the treatment group displays some special crime dynamics over the (pseudo-)treatment period in the years before 2014. To this end, I drop 2014 and, for each year y from 2006 to 2012, I run Equation (III.1) as if the World Cup and CPCopa program had taken place in such year. I do not perform this test for 2013 because the FIFA Confederations Cup (a smaller football tournament acting as a rehearsal for the World Cup) took place in Brazil

that year roughly over the pseudo-treatment period, and this may affect the test's results. I plot the pseudo-treatment coefficients with their 95% Conley and Taber confidence intervals in Figure III.A-3, together with the actual treatment coefficients for year 2014. The results confirm that the special dynamic in total crime and robbery affecting the treatment group is specific to 2014.

As a second placebo test, I exploit the discontinuity represented by the start of the CPCopa program. I drop the period May 20th 2014 to July 20th 2014 (the CPCopa period) and re-estimate my baseline regression using, as pseudo-treatment, the period before the CPCopa deployment. To this end, I modify equation (III.1) as follows:

(III.2) $Crime_{s,t} = \beta_0 + \beta_1(bf\_CPCopa_t * (Stadium_s + TC_s)) + \beta_2 bf\_CPCopa_t + \beta_5 ArenaSP_t + Z_t + \rho_t + \theta_s + \varepsilon_{st}$

Where bf_CPCopa_t is a dummy equal to one for those days in 2014 before the extra monitoring at the stadium started (January 1st to May 19th).

This placebo test allows me to check whether the stadium and TC PDs were exhibiting a different crime dynamic with respect to non-target PDs before the police increase. In such case, perhaps I am capturing a spurious correlation rather than a causal impact. Results are displayed in Table III.A-7. Using my preferred inference approach (Conley and Taber), outcomes validate my exercise in that they reveal no special crime dynamic affecting the treatment group before the CPCopa period.

The increase in police presence during the program period may have pushed criminals to put off their illegal activities to later times, causing crime to rise in the treated PDs after the end of the CPCopa period. To check for medium-term temporal displacement, I perform an exercise analogous to the one above using, as pseudo-treatment, the period after the CPCopa activity (July 21st 2014 to December 31st 2014). The results shown in Table III.A-7 suggest that there was no medium-term temporal displacement. This in turn implies that the crimes prevented by increased police presence represent averted – rather displaced – offenses.

In my baseline analysis I exclude from the sample all game days during the CPCopa period, in order to neutralize the voluntary incapacitation effect. I show in Panel (B) in Table III.B-1 that my baseline findings on the effect of police on crime are robust to an alternative specification where I include and control for World Cup game days. Still, one may worry that if a short-term temporal displacement effect was at play, whereby criminals diverted their illegal activities to

game days, then my results may be biased. Neglecting such effect would generate overly large (in absolute value) estimates, as it would inflate the crime reduction in treated PDs. To assess this concern, Table III.A-8 reports the average number of crimes in game days, non-game days during the CPCopa period and remaining days in 2014 in the stadium, TC and non-target PDs. In line with the previsions for the voluntary incapacitation effect, the average number of crimes during game days was slightly lower with respect to the rest of the CPCopa period. This descriptive evidence suggests that crime was not diverted to game days.

III.7 Conclusions

In this paper I provide novel, robust evidence on the causal impact of police on crime. I use the natural experiment represented by the creation of a special policing unit to monitor a few tournament-related areas in São Paulo during the 2014 FIFA World Cup. As an innovation with respect to previous studies, I employ a conceptual framework that is intended to account for the different mechanisms through which the World Cup may affect local crime. This approach allows me to root out alternative channels and isolate the specific impact of interest.

The DID estimates show that a police increase leads to a significant reduction in criminal activity. The total number of offenses per day decreases by 18%, and the number of robberies per day drops by 34%. There is no evidence of spatial (except for thefts) or temporal displacement, so that the measured crime reductions in treated districts actually imply prevented – rather than displaced – offenses. My results can be best interpreted as reflecting the deterrence effect, by which increased police presence reduces crime by making it more costly for potential offenders. I test the robustness of my empirical strategy by performing a number of placebo regressions. Outcomes suggest that I am indeed capturing the causal impact of police on crime rather than some spurious correlation deriving from different crime dynamics in the treatment and the control groups.

My estimate of the elasticity of crime to policing is remarkably close to the estimates obtained in previous studies, which looked at different natural experiments in different contexts. This suggests that a -0.3 to -0.4 elasticity may be considered as a reasonable approximation of the impact of police on crime.

Appendix III.A Figures and tables

TABLE III.A-1: CPCOPA MONITORED SPOTS AND NUMBER OF PDS WHERE THEY ARE LOCATED

Monitored spots					PDs		
Stadium	FIFA Hotels	Training centers	Venues of WC-related events	Transport and Touristic	Target	Non-target	Total
1	3	3	3	6	11	82	93

Note: Different monitored spots may be located in the same PD.

FIGURE III.A-1: THE TIMELINE OF THE WORLD CUP AND OF THE CPCOPA PROGRAM

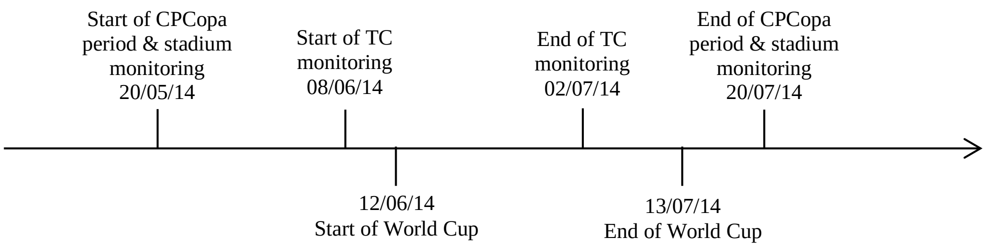


TABLE III.A-2: THE EXPECTED EFFECTS OF THE WORLD CUP ON LOCAL CRIME

Mechanism	Expected impact on crime
Police increase	⇓
Voluntary incapacitation	⇓
Concentration	⇑

Note: Upward and downward pointing arrows represent respectively a positive and negative effect from each of the three mechanisms – police increase, incapacitation and concentration – through which the World Cup may impact local crime.

TABLE III.A-3: CRIME DATA DAILY BY PD, 2006-2014 – DESCRIPTIVE STATISTICS

Statistic	Total crime	Robbery	Theft	Assault	Murder	Rape	Drugs
Mean	11.1	4.5	5.2	1.1	0.1	0.1	0.2
Min	0	0	0	0	0.00	0	0
Obs = 0 (%)	0.2	2.8	5.3	37.6	92.4	94.5	86.2
Max	211	132	147	15	5.00	5	10
Variance	37.2	10.0	15.4	1.5	0.1	0.1	0.2
Skewness	1.7	1.4	2.6	1.4	3.9	4.5	3.3
N	305,505	305,505	305,505	305,505	305,505	305,505	305,505

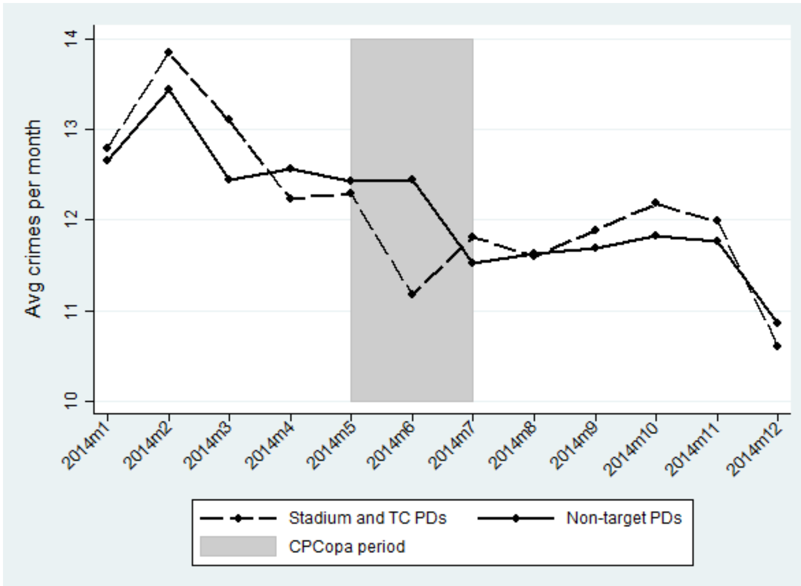
Note: Summary statistics are generated from the crimes reported daily in each of São Paulo's 93 PDs over the period 2006-2014. "Murder" comprises committed and attempted murders.

TABLE III.A-4: THE EXPECTED EFFECTS OF THE WORLD CUP ON LOCAL CRIME – DETAILED

Expected effect on crime	Non-target PDs		Target PDs			
			Stadium and TC		Others	
Game day	Yes	No	Yes	No	Yes	No
Police increase	-	-	⇓	⇓	⇓	⇓
Voluntary incapacitation	⇓	-	⇓	-	⇓	-
Concentration	-	-	⇑	-	⇑	⇑

Note: Upward and downward pointing arrows represent respectively a positive and negative effect from each of the three mechanisms – police increase, incapacitation and concentration – through which crime in target or non-target PDs may be impacted on game and non-game days during the World Cup. Dashes mean that no effect is expected.

FIGURE III.A-2: GRAPHIC REPRESENTATION OF THE DID APPROACH – AVERAGE NUMBER OF CRIMES PER MONTH ON NON-GAME DAYS IN THE STADIUM AND TC PDS *VS* NON-TARGET PDS, 2014



Notes: The averages are generated from the crimes reported monthly in the specified PDs in 2014, excluding days in which a FIFA World Cup game was played.

TABLE III.A-5: THE EFFECT OF POLICE ON CRIME – BASELINE RESULTS

Model	DV	$\widehat{\beta_1}$	95% Confidence interval		Inference method	N
			Lower	Upper		
Treatment group = stadium and TC PDs						
(A) Baseline time fixed effects	ln(Total crime)	-0.18*	-0.32	-0.06	Conley-Taber	273,402
	ln(Robbery)	-0.34*	-0.58	-0.15	Conley-Taber	258,323
	ln(Theft)	-0.08	-0.25	0.06	Conley-Taber	265,450
(B) Alternative time fixed effects	ln(Total crime)	-0.18*	-0.32	-0.06	Conley-Taber	273,402
	ln(Robbery)	-0.34*	-0.58	-0.15	Conley-Taber	258,323
	ln(Theft)	-0.08	-0.25	0.06	Conley-Taber	265,450
Treatment group = all target PDs						
(C) Baseline time fixed effects	ln(Total crime)	-0.08	-0.16	0.01	Clustered by PD	302,739
	ln(Robbery)	-0.18*	-0.29	-0.07	Clustered by PD	287,080
	ln(Theft)	0.01	-0.08	0.09	Clustered by PD	294,678

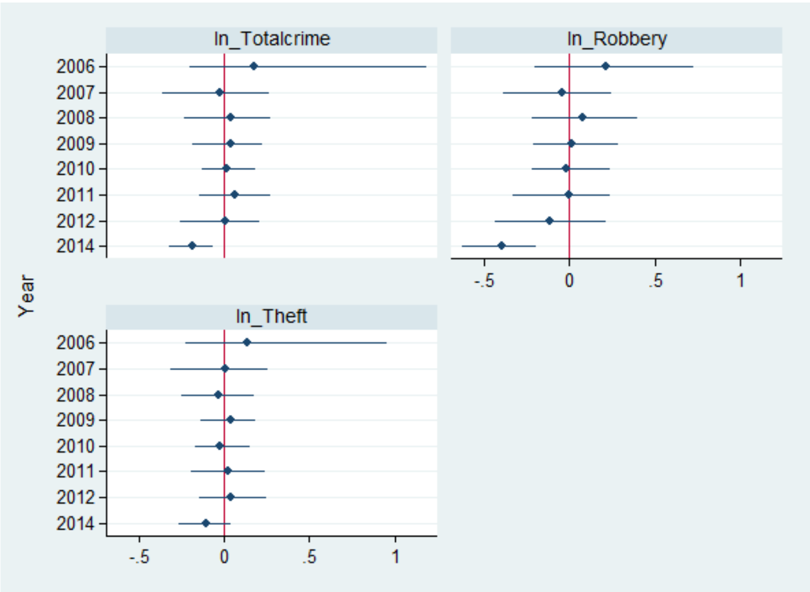
*Notes: An observation is a day-long period for one of the considered PDs in São Paulo over the period 2006-2014, excluding days in which a 2014 FIFA World Cup game was played. Considered PDs in panels (A) and (B) comprise 84 districts (the stadium, TC, and non-target ones). Considered PDs in panel (C) comprise all 93 districts in São Paulo. The estimates come from ordinary least squares regressions with PD and time fixed effects. Baseline time fixed effects include indicators for the year, month, day of the month and day of the week. Alternative time fixed effects include indicators for the year, day of the year and day of the week. Other controls include a dummy for the CPCopa period, a dummy for the period over which the TC received extra monitoring, an indicator for the World Cup period, an indicator for the period after the inauguration of the São Paulo Arena, a holiday indicator and a dummy for the period when it was possible to file robbery reports online. * = significantly different from zero at the 5% level using the specified inference method.*

TABLE III.A-6: SPATIAL DISPLACEMENT AND SPILL-OVER EFFECT

Analysis	Specification	DV	$\widehat{\beta_1}$	95% Confidence interval		Inference method	N
				Lower	Upper		
(A) Spatial displacement	T = stadium and TC PDs C = adjacent PDs	ln(Total crime)	-0.19*	-0.23	-0.13	Conley-Taber	42,355
		ln(Robbery)	-0.33*	-0.42	-0.21	Conley-Taber	40,720
		ln(Theft)	-0.12*	-0.18	-0.08	Conley-Taber	41,786
(B) Spill-overs	T = adjacent PDs C = other non-target PDs	ln(Total crime)	0.04	-0.05	0.13	Clustered by PD	266,887
		ln(Robbery)	-0.02	-0.18	0.14	Clustered by PD	252,129
		ln(Theft)	0.10*	0.02	0.18	Clustered by PD	259,023

*Notes: T = Treatment group; C = Control group. An observation is a day-long period for one of the considered PDs in São Paulo over the period 2006-2014, excluding days in which a 2014 FIFA World Cup game was played. Considered PDs in panel (A) comprise 13 districts. Considered PDs in panel (B) comprise all 82 non-target districts. The estimates come from ordinary least squares regressions with PD and time fixed effects. Time fixed effects include indicators for the year, month, day of the month and day of the week. Other controls include a dummy for the CPCopa period, a dummy for the period over which the TC received extra monitoring, an indicator for the World Cup period, an indicator for the period after the inauguration of the São Paulo Arena, a holiday indicator and a dummy for the period when it was possible to file robbery reports online. * = significantly different from zero at the 5% level using the specified inference method.*

FIGURE III.A-3: PSEUDO-TREATMENT COEFFICIENTS (2006-2012) AND ACTUAL TREATMENT COEFFICIENTS (2014)



Notes: For each DV (total crime, robbery and theft and for each year y over the period 2006-2012, the figures report the estimated pseudo-treatment coefficients and 95% confidence intervals from running Equation (III.1) as if the World Cup and CPCopa program had taken place in year y. The figures also report the actual treatment coefficient and confidence interval (for year 2014) derived from the baseline analysis. I do not perform the test for year 2013 because the FIFA Confederations Cup place in Brazil that year roughly over the pseudo-treatment period. All confidence intervals are calculated using the Conley and Taber inference approach.

TABLE III.A-7: CRIME DYNAMICS IN THE TREATED PDS IN THE PERIODS SURROUNDING THE CPCOPA PROGRAM

Analysis	Pseudo-treatment period	DV	$\widehat{\beta_1}$	95% Confidence interval		Inference method	N
				Lower	Upper		
(A) Placebo	01/01/14-19/05/14	ln(Total crime)	-0.16	-0.35	0.02	Conley-Taber	270,295
		ln(Robbery)	-0.25	-0.52	0.07	Conley-Taber	255,298
		ln(Theft)	-0.07	-0.24	0.07	Conley-Taber	262,407
(B) Temporal displacement	21/07/14-31/12/14	ln(Total crime)	-0.18	-0.41	0.04	Conley-Taber	270,295
		ln(Robbery)	-0.15	-0.46	0.19	Conley-Taber	255,298
		ln(Theft)	-0.16*	-0.35	-0.01	Conley-Taber	262,407

*Notes: An observation is a day-long period for one of the considered PDs in São Paulo over the period 2006-2014, excluding the CPCopa period. Considered PDs comprise 84 districts (the stadium, TC, and non-target ones). The estimates come from ordinary least squares regressions with PD and time fixed effects. Time fixed effects include indicators for the year, month, day of the month and day of the week. Other controls include a dummy for the pseudo-treatment period, an indicator for the period after the inauguration of the São Paulo Arena, a holiday indicator and a dummy for the period when it was possible to file robbery reports online. * = significantly different from zero at the 5% level using the specified inference method.*

TABLE III.A-8: AVERAGE NUMBER OF CRIMES ON GAME DAYS, NON-GAME DAYS DURING THE CPCOPA PERIOD AND ALL OTHER DAYS IN 2014

Period	Non-target PDs (average)	Stadium PD	TC PD
2014 excl. CPCopa period	12.1	7.4	17.1
CPCopa period excl. game days	12.1	7.0	14.8
Game days	11.3	5.9	14.6

Notes: Means are generated from the crimes reported daily in the specified PDs over the specified days. For non-target and stadium PDs, the CPCopa period is defined as May 20th 2014 to July 20th 2014. For the TC PD, the CPCopa period is defined as June 8th 2014 to July 2nd 2014. In the stadium PD case, I exclude the six days in which a World Cup game was played at the São Paulo Arena, as criminal activity was exceptionally high (likely due to concentration).

Appendix III.B Additional tables

TABLE III.B-1: THE EFFECT OF POLICE ON CRIME – ALTERNATIVE MODELS

Model	DV	$\widehat{\beta_1}$	95% Confidence interval		Inference method	N
			Lower	Upper		
(A) Controlling for non-treated target PDs	ln(Total crime)	-0.18*	-0.33	-0.06	Conley-Taber	302,739
	ln(Robbery)	-0.34*	-0.58	-0.15	Conley-Taber	287,080
	ln(Theft)	-0.08	-0.29	0.06	Conley-Taber	294,678
(B) Controlling for game days	ln(Total crime)	-0.23*	-0.29	-0.18	Conley-Taber	275,499
	ln(Robbery)	-0.39*	-0.44	-0.34	Conley-Taber	260,359
	ln(Theft)	-0.10*	-0.15	-0.06	Conley-Taber	267,497

*Notes: An observation is a day-long period for one of the considered PDs in São Paulo over the considered period. The analysis in panel (A) considers all 93 PDs in São Paulo over the period 2006-2014, excluding days in which a 2014 FIFA World Cup game was played. The analysis in panel (B) considers 84 PDs (the stadium, TC, and non-target ones) over the period 2006-2014. The estimates come from ordinary least squares regressions with PD and time fixed effects. Time fixed effects include indicators for the year, month, day of the month and day of the week. All regressions include a dummy for the CPCopa period, a dummy for the period over which the TC received extra monitoring, an indicator for the World Cup period, an indicator for the period after the inauguration of the São Paulo Arena, a holiday indicator and a dummy for the period when it was possible to file robbery reports online. In addition, regression in panel (A) includes a dummy for the interaction between non-treated target PDs and the CPCopa period; regression in panel (B) includes indicators for game days, days when a match was played in São Paulo and a dummy for the interaction between the latter variable and the stadium PD. * = significantly different from zero at the 5% level using the specified inference method.*

TABLE III.B-2: THE EFFECT OF POLICE ON CRIME – DEPENDENT VARIABLE IN LEVELS

DV	$\widehat{\beta_1}$	95% Confidence interval		Inference method	N
		Lower	Upper		
Total crime	-2.08*	-3.90	-0.81	Conley-Taber	273,840
Robbery	-1.43*	-2.86	-0.39	Conley-Taber	273,840
Assault	-0.61	-1.35	0.14	Conley-Taber	273,840
Rape	-0.14	-0.35	0.10	Conley-Taber	273,840
Murder	-0.04	-0.11	0.00	Conley-Taber	273,840
Theft	-0.02	-0.07	0.05	Conley-Taber	273,840
Drugs	0.16	-0.04	0.24	Conley-Taber	273,840

*Notes: An observation is a day-long period for one of the considered PDs in São Paulo over the period 2006-2014, excluding days in which a 2014 FIFA World Cup game was played. Considered PDs comprise 84 districts (the stadium, TC, and non-target ones). The estimates come from ordinary least squares regressions with PD and time fixed effects. Time fixed effects include indicators for the year, month, day of the month and day of the week. Other controls include a dummy for the CPCopa period, a dummy for the period over which the TC received extra monitoring, an indicator for the World Cup period, an indicator for the period after the inauguration of the São Paulo Arena, a holiday indicator and a dummy for the period when it was possible to file robbery reports online. "Murder" includes committed and attempted murders. * = significantly different from zero at the 5% level using the specified inference method.*

References

Anderson C.A., Shibuya, A., Ihori, N., Swing, E.L., Bushman, B.J., Sakamoto, A., Rothstein, H.R. and Saleem, M. (2010) "Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review." *Psychological Bulletin*, Vol. 136(2): 151–73.

Anderson, C.A., Gentile, D.A. and Buckley, K.E. *Violent video game effects on children and adolescents: theory, research and public policy*. New York: Oxford University Press, 2007.

BBC, "Brasil lidera ranking de medo de tortura policial." May 12th 2014. Available: http://www.bbc.com/portuguese/noticias/2014/05/140512_brasil_tortura_vale_rb. Accessed: March 31st 2017.

Becker, G. (1968) "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, Vol. 76(2): 169–217.

Bellou, A. (2015) "The Impact of Internet Diffusion on Marriage Rates: Evidence from the Broadband Market." *Journal of Population Economics*, Vol. 28: 265–297.

Bhuller, M., Havnes, T., Leuven, E. and Mogstad, M. (2013) "Broadband Internet: An Information Superhighway to Sex Crime?" *Review of Economic Studies*, Vol. 80(4): 1237–1266.

Campaniello, N. (2013) "Mega events in sports and crime: evidence from the 1990 Football World Cup." *Journal of Sports Economics*, Vol. 14: 148–170.

Card, D. and Dahl, G. (2011) "Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior." *The Quarterly Journal of Economics*, Vol. 126: 103–43.

Conley, T.G. and Taber, C.R. (2011) "Inference with 'Difference in Differences' with a Small Number of Policy Changes." *The Review of Economics and Statistics*, Vol. 93 (1): 113-125.

Cunningham, S., Engelstätter, B. and Ward, M.R. (2016), "Violent Video Games and Violent Crime." *Southern Economic Journal*, Vol. 82: 1247–1265.

Dahl, G. and Della Vigna, S. (2009) "Does movie violence increase violent crime?" *The Quarterly Journal of Economics*, Vol. 124(2): 677–734.

Di Tella, E. and Schargrodsky, E. (2004) "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces after a Terrorist Attack." *American Economic Review*, Vol. 94: 115–33.

DiIulio, J. (1996) "Help Wanted: Economists, Crime, and Public Policy." *Journal of Economic Perspectives*, Vol. 10: 3–24.

Donohue, J. and Levitt, S. (2001) "Legalized Abortion and Crime." *Quarterly Journal of Economics*, Vol. 116(2): 379–420.

Draca, M., Machin, S. and Witt, R. "Crime Displacement and Police Interventions: Evidence from London's 'Operation Theseus'." In *The Economics of Crime: Lessons for and from Latin America*, Ed. Rafael Di Tella, Sebastian Edwards and Ernesto Schargrodsky. Chicago MA: University of Chicago Press, 2010, p. 359–78.

Draca, M., Machin, S. and Witt, R. (2011) "Panic on the streets of London: police, crime and the July 2005 terror attack." *American Economic Review*, Vol. 101: 2157–2181.

Dworkin, A. *"Pornography": Men Possessing Women*. London: The Women's press, 1981.

Ehrlich, I. (1981) "On the Usefulness of Controlling Individuals: An Economic Analysis of Rehabilitation, Incapacitation and Deterrence." *American Economic Review*, Vol. 71: 307-332.

FBI – Internet Crime Complaint Center (IC3), *2013 Internet crime report*. Available: http://www.ic3.gov/media/annualreport/2013_IC3Report.pdf. Accessed: March 31st 2017.

FBI, *Crime in the United States 2011*. Available: https://ucr.fbi.gov/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/offense-definitions. Accessed: March 28th 2017.

FBI, *Crime in the United States 2013*. Available: https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013. Accessed: March 31st 2017.

FBI, *UCR Frequently Asked Questions*. Available: https://www2.fbi.gov/ucr/ucr_general.html. Accessed: March 28th 2017.

FCC, *The Broadband Availability Gap, 2010*. Available: https://transition.fcc.gov/national-broadband-plan/broadband-availability-gap-paper.pdf. Accessed: March 31st 2017.

FIFA, "2014 FIFA World Cup™ reached 3.2 billion viewers, one billion watched final." December 16th 2015. Available: http://www.fifa.com/worldcup/news/y=2015/m=12/news=2014-fifa-world-cuptm-reached-3-2-billion-viewers-one-billion-watched--2745519.html. Accessed: March 31st 2017.

FIFA, *Regulations. 2014 FIFA World Cup Brazil*. Available: http://www.fifa.com/mm/document/tournament/competition/01/47/38/17/regulationsfwcbrazil2014_update_e_neutral.pdf. Accessed: March 31st 2017.

Folha de Jardim, "40 pontos estratégicos de SP serão vigiados por 4 mil PMs durante a Copa." May 16th 2014. Available: http://www.folhadejardim.com.br/40-pontos-estrategicos-de-sp-serao-vigiados-por-4-mil-pms-durante-a-copa/. Accessed: March 31st 2017.

Folha de S. Paulo, "Mortes por policiais no país em um ano equivalem a um 11 de Setembro." October 3rd 2015. Available: http://www1.folha.uol.com.br/cotidiano/2015/10/1689789-mortes-por-policiais-no-pais-em-um-ano-equivalem-a-um-11-de-setembro.shtml. Accessed: March 31st 2017.

Folha de S. Paulo, "SP aumentará efetivo de policiais militares." June 3rd 2014. Available: http://www1.folha.uol.com.br/poder/2014/06/1464086-sp-aumentara-efetivo-de-policiais-militares.shtml. Accessed: March 31st 2017.

Fox, J. *Trends in Juvenile Violence: A Report to the United States Attorney General on Current and Future Rates of Juvenile Offending*. Washington, D.C.: Bureau of Justice Statistics, 1996.

Griffiths, M. and Sutton, M. (2013) "Proposing the Crime Substitution Hypothesis: Exploring the possible causal relationship between excessive adolescent video game playing, social networking and crime reduction." *Education and Health Journal*, Vol. 31(1): 17–21.

ITU "The Impact of Broadband on the Economy: Research to Date and Policy Issues." April 2012. Available: https://www.itu.int/ITU-D/treg/broadband/ITU-BB-Reports_Impact-of-Broadband-on-the-Economy.pdf. Accessed: March 31st 2017.

Kalist, D.E., and Lee, D.Y. (2016) "The National Football League: Does Crime Increase on Game Day?" *Journal of Sports Economics*, Vol. 17(8): 863–882.

Kelly, M. (2000) "Inequality And Crime." *The Review of Economics and Statistics*, Vol. 82(4): 530–539.

Kendall, T.D. (2007) "Pornography, rape, and the Internet." Mimeo, Clemson University.

King, M., Ruggles, S., and Trent, A., Flood, S., Genadek, K., Schroeder ,M. B., Trampe, B., and Vick, R. (2010) Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota.

Kirby, S., Francis, B. and O'Flaherty, R. (2014) "Can the FIFA World Cup Football (Soccer) Tournament be Associated with an Increase in Domestic Abuse?" *Journal of Research in Crime and Delinquency*, Vol. 51(3): 259–276.

Klick, J. and Tabarrok, A. (2005) "Using terror alert levels to estimate the effect of police on crime." *Journal of Law and Economics*, Vol. 68: 267–279.

Kuhn, P. and Mansour, H. (2014) "Is Internet Job Search Still Ineffective?" *The Economic Journal*, Vol. 124: 1213–1233.

Laqueur, H. and Copus, R. "Entertainment as Crime Prevention: Evidence from Chicago Sports Games" (April 25, 2014). Available at SSRN: https://ssrn.com/abstract=2429551. Accessed: March 31st 2017.

Lehr, W.H., Osorio, C.A., Gillett, S.E. and Sirbu, M.A. (2006). "Measuring broadband's economic impact." Final Report. National Technical Assistance, Training, Research and Evaluation Project # 99-07-13829. Available: http://cfp.mit.edu/publications/CFP_Papers/Measuring_bb_econ_impact-final.pdf. Accessed: March 31st 2017.

Levitt, S.D. (2002) "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply." *American Economic Review*, Vol. 92: 1244–50.

Levitt, S.D. (1997) "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *American Economic Review*, Vol. 87(3): 270–90.

Levitt, S.D. (2004) "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not." *Journal of Economic Perspectives*, Vol. 18(1): 163–190.

Mackinnon, C. *Only Words*. London: Harper Collins, 1995.

Marie, O. (2016) "Police and thieves in the stadium: measuring the (multiple) effects of football matches on crime." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 179: 273–292.

McCrary, J. (2002) "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment." *American Economic Review* Vol. 92: 1236–43.

NTIA, *A nation online: entering the broadband age*. U.S. Department of Commerce, Economics and Statistics Administration, 2004.

NTIA, *Exploring the digital nation: home broadband Internet adoption in the United States*. U.S. Department of Commerce, Economics and Statistics Administration, 2011.

NTIA, "Household Broadband Adoption Climbs to 72.4 Percent." June 6th 2013. Available: http://www.ntia.doc.gov/blog/2013/household-broadband-adoption-climbs-724-percent. Accessed: March 31st 2017.

Pew Research Center, "One-fifth of Americans report going online 'almost constantly'." December 8th 2015. Available: http://www.pewresearch.org/fact-tank/2015/12/08/one-fifth-of-americans-report-going-online-almost-constantly/. Accessed: March 31st 2017.

Portal da Copa - Brazilian Government official website on the 2014 World Cup. Available: http://www.copa2014.gov.br/pt-br/noticias. Accessed: March 31st 2017.

Posner, R.A. *Sex and Reason*. Cambridge, MA: Harvard University Press, 1994.

Poutvaara, P. and Priks, M. (2009) "The effect of police intelligence on group violence: evidence from reassignments in Sweden." *Journal of Public Economics*, Vol. 93: 403–411.

Prefeitura de São Paulo, "Estádio da abertura da Copa 2014 como dinamizador do desenvolvimento da Zona Leste e da Cidade de São Paulo." January 2011. Available: http://www.prefeitura.sp.gov.br/cidade/secretarias/upload/trabalho/arquivos/itaquera.pdf. Accessed: Accessed: March 31st 2017.

Prefeitura de São Paulo, "Balanço das ações da Copa do Mundo em São Paulo." July 11th 2014. Available: http://capital.sp.gov.br/noticia/balanco-das-acoes-da-copa-do-mundo-em-sao-paulo. Accessed: March 31st 2017.

Rees, D.I., and Schnepel, K.T. (2009) "College Football Games and Crime." *Journal of Sports Economics*, Vol. 10(1): 68–87.

Stevenson, B. (2006) "The impact of the Internet on worker flows". Working paper. The Wharton School, University of Pennsylvania.

United States of America. Telecommunications Act 1996. Available: http://www.internet.org.za/telecoms_act.html. Accessed: March 31st 2017.

The Economist, "Police violence in Brazil. Serial killing." March 20th 2014. Available: http://www.economist.com/blogs/americasview/2014/03/police-violence-brazil. Accessed: March 31st 2017.

The World Bank, List of countries by intentional homicide rate. Available: http://data.worldbank.org/indicator/VC.IHR.PSRC.P5?order=wbapi_data_value_2013+wbapi_data_value+wbapi_data_value-last&sort=desc&year_high_desc=true. Accessed: March 31st 2017.

Wallsten, S. (2013) "What Are We Not Doing When We're Online." NBER Working Paper No. 19549. Available: http://www.nber.org/papers/w19549. Accessed: March 31st 2017.

Zillman, D. (1971) "Excitation transfer in communication-mediated aggressive behavior." *Journal of Experimental Social Psychology*, Vol. 7: 419–434.

Zimring, F.E. and Hawkins, G. *Crime Is Not the Problem: Lethal Violence in America*. New York: Oxford University Press, 1997.