Pontifícia Universidade Católica do Rio de Janeiro

Bernardo Rennó Duque

Who Becomes a Criminal? Evidence From Brazil

Dissertação de Mestrado

Master's Dissertation presented to the Programa de Pós–graduação em Economia, do Departamento de Economia da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia.

Advisor : Prof. Gustavo Gonzaga
Co-advisor: Prof. Claudio Ferraz

Rio de Janeiro
April 2024

Pontifícia Universidade Católica do Rio de Janeiro

Bernardo Rennó Duque

Who Becomes a Criminal? Evidence From Brazil

Master's Dissertation presented to the Programa de Pós–graduação em Economia da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia. Approved by the Examination Committee:

Prof. Gustavo Gonzaga
Advisor
Departamento de Economia – PUC-Rio

Prof. Claudio Ferraz
Co-advisor
UBC & PUC-Rio

Prof. Tomás Guanziroli
PUC-Rio

Prof. Joana Monteiro
FGV EBAPE

Rio de Janeiro, April 26th, 2024

Bernardo Rennó Duque

B.A. in Economics, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), 2020

Bibliographic data

Duque , Bernardo Rennó

Who Becomes a Criminal? Evidence From Brazil / Bernardo Rennó Duque; advisor: Gustavo Gonzaga; co-advisor: Claudio Ferraz. – 2024.

40 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, 2024.

Inclui bibliografia

1. Economia – Teses. 2. Crime. 3. Criminosos. 4. Criminalidade. 5. Decisão Criminal. I. Gonzaga, Gustavo. II. Ferraz, Claudio. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 330

Acknowledgments

First, I would like to thank my parents, Claudia and Moises. I could not have gotten here without your love and support. To my mother, thank you for always trusting in my potential and allowing me to buy as many books as I wanted. To my father, thank you for nagging me to study during my teenage years when all I wanted was to play video-games.

I am very grateful to my grandparents, Clayton and Sandra, for always supporting and providing me with everything I needed and more. I have always admired both of you for your intelligence. The stories you told me will always be with me. I am also very grateful to the rest of my family for all their love, and to my eternal professor, Luis Cesar.

Many thanks to the friends I made at PUC during my Master's. You surely made things much easier and less lonely. Thanks to my other lifelong friends, who have always been there for me and have given me a different perspective on how to see the world.

Finally, I would like to thank the committee members, Tomás and Joana, for all their comments, as well as my official and unofficial advisors, Breno, Claudio, Diogo, Gustavo and Ricardo.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and the Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq.

Abstract

Duque , Bernardo Rennó; Gonzaga, Gustavo (Advisor); Ferraz, Claudio (Co-Advisor). **Who Becomes a Criminal? Evidence From Brazil**. Rio de Janeiro, 2024. 40p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

We investigate selection into criminality. Specifically, we use a novel dataset on inmates across 24 of 26 Brazilian states and rich administrative data to draw the profiles of criminals born between 1996 and 2002. We compare them with non-criminal individuals from the same cohort using a linear probability model with year of birth and place fixed effects. Our findings suggest that criminals come from more urban and vulnerable places, with a higher exposure to other criminals. In addition, our results indicate that family structure matters and that parents of criminals are more economically vulnerable than parents of non-criminals. Finally, criminals have worse educational performance, a history of juvenile detention, and worse labor market performance.

Keywords

Crime; Criminals; Criminality; Criminal Decision.

Resumo

Duque , Bernardo Rennó; Gonzaga, Gustavo; Ferraz, Claudio. **Quem Se Torna um Criminoso? Evidência do Brasil**. Rio de Janeiro, 2024. 40p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Nós investigamos seleção à criminalidade. Especificamente, nós usamos uma base de dados inédita sobre prisioneiros de 24 dos 26 estados brasileiros e ricos dados administrativos para traçar os perfis de criminosos nascidos entre 1996 e 2002. Nós os comparamos com não-criminosos da mesma coorte usando um modelo de probabilidade linear com efeitos fixos de ano de nascimento e de lugar. Nossos achados sugerem que criminosos vem de lugares mais urbanos e vulneráveis, com uma maior exposição a outros criminosos. Além disso, nossos resultados indicam que a estrutura familiar importa e que pais de criminosos são economicamente mais vulneráveis do que pais de não-criminosos. Por fim, criminosos tem piores resultados educacionais, um histórico de detenção juvenil, e piores resultados no mercado de trabalho.

Palavras-chave

Crime; Criminosos; Criminalidade; Decisão Criminal.

Table of contents

List of figures

List of tables

List of Abreviations

CAF - Development Bank of Latin America and the Caribbean
CPF - Cadastro de Pessoas Físicas
DEPEN - Departamento Penitenciário Nacional
DF - Federal District
ECA - Estatuto da Criança e do Adolescente
INFOPEN - Levantamento Nacional de Informações Penitenciárias
RAIS - Relação Anual de Informações Sociais
SAEB - Sistema de Avaliação da Educação Básica
UNODC - United Nations Office on Drugs and Crime

1
Introduction

Crime bears both direct and indirect costs to societies. Direct costs include, for instance, lower life expectancy (Anderson, 1999) and expenses for public and private security (Becker, 1968), while indirect costs encompass worse school performance (Monteiro & Rocha, 2017), loss of working days (Anderson, 1999), and even more future crime (Damm & Dustmann, 2014; Sviatschi, 2022). Due to the illegal nature of criminal operations and the consequent scarcity of data on the subject, evidence on who becomes a criminal is limited. Understanding the profile of offenders is key to formulating public policies designed to deter individuals from pursuing this path. By implementing targeted interventions that address this profile, it may be possible to reduce the prevalence of illegal behavior and increase the likelihood of positive outcomes for those at risk of engaging in such activities.

In this paper, we examine the main selection patterns of individuals who follow a criminal path. To do this, we use a novel dataset on inmates from 24 out of 26 Brazilian states and the Federal District — provided by the Brazilian Ministry of Justice — to create a "registry of criminals". We merge this registry with various administrative datasets to document the profile of criminals, using the general population as a comparison group. The richness of the data allows us to identify not only criminals and non-criminals but also their parents and the neighborhoods where these individuals grew up. These features enable us to provide, to the best of our knowledge, the first comprehensive analysis of this kind.

Our study is descriptive; we are interested in characterizing the differences between criminals and non-criminals. Previous literature has focused on developed countries (e.g., Laub & Sampson, 1993; Sampson & Laub, 2005; Wikström, 2009; Levitt & Venkatesh, 2001). We add to this body of work by providing evidence for a developing country characterized by high crime rates and a strong perception of insecurity. Our analysis focuses on the cohort born between 1996 and 2002, using linear probability models with year-of-birth and place fixed effects to generate five sets of results.

First, we find that criminals come from different neighborhoods compared to non-criminals. They are more likely to originate from vulnerable and urban areas, with higher exposure to other criminals. Specifically, in our preferred specification, we find that, all else equal, an individual born in a Brazilian slum in an urban area has a 46.7% higher chance of becoming a criminal than someone born in a rural area outside a slum.[1] Additionally, a 10 percentage points increase in the share of criminals in a neighborhood is associated with a 3.6% increase in the probability of an individual from that area following a criminal path. This is consistent with peer or role model effects. These results reinforce the need to control for place fixed effects, which is something we do in the rest of the analysis.

Second, we find that family structure matters. Individuals have a 9.6% lower chance of becoming a criminal if their father is present. In addition, if a mother gives birth at the age of 20 instead of 16, all else equal, this is associated with a 14.7% decrease in the probability of her child choosing a criminal path. Finally, an individual born into a family with four siblings has a 10% higher chance of becoming a criminal compared to someone from a family with only one child.

Third, our results indicate that criminals come from more economically vulnerable families. Parents who receive welfare benefits when their child is under 19 are associated with a 16.1% to 33.2% higher chance of future criminal behavior by their child. Moreover, parents who have a formal job when their child is between 0 and 10 are linked to a reduction in the probability of this child committing a crime during adulthood, ranging from 2.3% to 13.2%. Although we find a negative correlation between parental labor income and criminal activity, the magnitude of the coefficients is relatively small, indicating it may not be a significant predictor.

Fourth, criminals have been in juvenile detention more often and have worse educational performance than non-criminals. We find that juvenile detention is one of the most important predictors of criminal behavior in early adulthood. Specifically, it is associated with a 31% increase in the probability of criminal behavior between the ages of 18 and 23. Additionally, criminals exhibit higher age-grade distortion, linked to an 11.9% increase in the probability of criminal activity, and perform worse in national standardized exams in both 5th and 9th grade. However, the latter is a more important predictor than the former.[2] They also have a higher probability of attending public schools (22%), which is strongly correlated with worse educational opportunities in Brazil.

Finally, while the evidence is weaker, we find that criminals have worse labor market performance. They have lower formal labor income and shorter job spell durations, but the magnitudes of these variables are not significant. In contrast, they have significantly less stable jobs (−21.9%), which we measure

[1] 29.5% for urban area + 17.1% for slums.

[2] A 1 standard-deviation increase in the 9th grade exam score is associated with a 15.9% decrease in the chance of becoming a criminal, while the equivalent for the 5th grade is −4.8%.

using public jobs as a proxy. Moreover, each additional year of schooling is associated with a 5.6% decrease in the probability of becoming a criminal. Lastly, having a job is linked to a 22.8% reduction in criminal behavior, consistent with the literature, although this result is not stable across all specifications.

Related Literature. Our paper contributes to the literature on the decision to commit crimes, which has been explored by both Economics and Criminology. In Economics, this literature started with the seminal paper by Becker (1968), the first to propose an economic approach to the decision to follow a criminal path, and the extension made by Ehrlich (1973). There are also other studies that extend their idea of crime as an occupational choice by, for instance, modeling the decision to become a criminal over the life cycle (Fella & Gallipoli, 2014) and modeling the possibility of multiple types of criminal activities (Koskela & Virén, 1997).

Other studies have also expanded this literature by identifying channels that can lead to criminal involvement.[3] For instance, some papers stress the role of having a job and family ties on criminal activity: Britto, Pinotti & Sampaio (2022) have investigated the increased likelihood of committing crimes after losing a job; Britto, Melo & Sampaio (2022) the effects of parental job loss on the likelihood of their children committing crimes; and Britto et al. (2022b) the increasing effects on crime due to financial demands of having a child. In addition, some studies show the importance of education to decrease criminal activity (Heckman et al., 2010; Barr & Gibbs, 2022; Lochner & Moretti, 2004).[4] Furthermore, the literature also highlights the importance of the neighborhoods where one grows up, which can expose them to violence and crime, increasing future criminal behavior (Sviatschi, 2022; Damm & Dustmann, 2014; Aizer, 2008; Chetty & Hendren, 2018).[5]

However, these empirical studies do not attempt to draw the profile of criminals; instead, they focus on these specific channels that can lead someone to commit a crime. The exceptions are Carvalho & Soares (2016) and Levitt & Venkatesh (2001), which investigate selection into gang membership.[6] Nevertheless, they differ from our work as we use administrative data and focus

[3]For a review of the relationship between economic incentives and crime, see Draca & Machin (2015).

[4]Although some papers show that some types of criminals such as white collar criminals and *mafiosi* can sometimes benefit from more education (Lochner (2004); Campaniello, Gray & Mastrobuoni (2016)).

[5]For a point of view of criminology, Sharkey (2017) reviews the evidence of exposure to violence affecting criminal behavior.

[6]Levitt & Venkatesh (2000) also presents some characteristics of the places where these gang members come from, but the focus of the paper is on the functioning of gangs rather than on their members.

on criminals in general rather than gang members. Therefore, we contribute to this literature by drawing a comprehensive profile of criminals in general.

Within the field of Criminology, the literature stresses that following a criminal path is not a permanent decision and that age is a very important predictor of criminal behavior. We highlight two theories that try to explain these facts, which can be seen as complementary. First, there is the Situational Action Theory (Wikström, 2009; CAF, 2014), which explains the commitment of crime as the interaction between crime propensity (e.g., self-control, moral values) and criminogenic exposure (e.g., peers, neighborhood environment), both subject to changes over time. The second theory is the Age-Graded Theory, which focuses on social bondings and turning points (Laub & Sampson, 1993; Sampson & Laub, 2003; Sampson & Laub, 2005).[7] The idea is that certain events in life, such as marriage and employment, can increase or decrease the social bonds that individuals have to society, which act as barriers to committing crimes. Both of these theories emphasize that the individuals' characteristics and environments should be considered together.[8]

Nevertheless, since all of these papers use data from developed countries and often suffer from small samples that lack external validity, their findings may not apply to nations with very different characteristics. Providing evidence for a developing country is important for designing effective, context-specific policies and interventions. Hence, we complement their research by testing some of their predictions using a large and representative sample of almost an entire cohort from a middle-income country characterized by high crime rates.

Finally, our study relates to three other papers that do not constitute a literature of their own but are somewhat related. These papers investigate the selection of individuals into various activities. Dal Bó et al. (2017), for instance, examine politicians in Sweden and find that they are positively selected in terms of ability while also being representative of the Swedish population with regard to demographics; Bell et al. (2018) investigate inventors and the influence of role models on encouraging children to pursue this path; and Levine & Rubinstein (2016) addresses entrepreneurs and their relation to high learning aptitude and high illicit activity scores. Our work adds to this body of research by investigating selection into criminality.

The remainder of this paper is organized as follows. Section 2 presents the context, the data we use, and some descriptives. Next, Section 3 presents

[7]For a review, see Nguyen & Loughran (2018).

[8]Also worth mentioning are other papers in Criminology that investigate adult criminal onset, underscoring the importance of outcomes both during and after childhood to understand criminal behavior (Mata & Dulmen, 2012; Zara & Farrington, 2009) and the negative effects parental imprisonment has on crime (Murray & Farrington, 2005).

the results, while Section 4 concludes.

2
Context and Data

2.1
Context

Brazil has one of the highest homicide rates in the world (UNODC, 2019), which, combined with the relatively high number of robberies and thefts, translates into a strong perception of insecurity.[1] Public security is also a major concern for voters, ranking second as a priority for policy.[2] This scenario is typically addressed not with crime prevention, but with a hardline approach. Such policies include police incursions into *favelas*,[3] which often result in casualties among civilians, police officers, and criminals, or a focus on arresting a high number of individuals, regardless of the severity of the crime committed and without aiming for rehabilitation.

This situation has made Brazil the country with the third-largest prison population in the world (Fair & Walmsley, 2021). The prison population has been doubling since 1990, when it was estimated to be around 90,000 people, and has now reached more than 811,000 inmates (INFOPEN, 2014; Fair & Walmsley, 2021).[4] Reports of human rights abuses and mistreatment of inmates are common in the media. Only individuals aged 18 or over can be arrested for committing crimes. If a child or teenager breaks the law, they are subject to the Statute of the Child and Adolescent (*Estatuto da Criança e do Adolescente, ECA*) instead of the penal code. As punishment for a crime, they can undergo socio-educational measures or be sent to juvenile detention.

These conditions contribute to the public's perception of criminals. Popular belief, subject to prejudices and stereotypes, suggests that criminals are typically low-income, low-educated individuals often coming from *favelas*. This perception likely arises because these areas are frequently home to various violent drug-trafficking organizations that control these regions through armed presence and exploitation of economic activities (Monteiro et al., 2022).

[1] See https://www1.folha.uol.com.br/cotidiano/2024/03/inseguranca-nas-ruas-a-noite-cresce-e-alcanca-2-de-cada-3-brasileiros-diz-datafolha.shtml

[2] See https://www1.folha.uol.com.br/poder/2023/12/datafolha-saude-e-principal-problema-do-pais-lula-derrapa-em-seguranca-e-corrupcao.shtml

[3] *Favelas* are Brazilian slums typically characterized by high-density population, low-quality infrastructure, and high levels of poverty.

[4] This represents around 381 people incarcerated per 100,000 citizens. Brazil ranks 15th in this comparison.

2.2 Data

2.2.1 Inmates

The Brazilian Ministry of Justice has provided us with the universe of inmates for 24 out of 26 Brazilian states and the Federal District (DF) since the early 2000s.[5] DEPEN (*Departamento Penitenciário Nacional*) is the public agency subordinated to the Ministry, which is responsible for compiling all data on inmates nationwide. Except for the federal prisons, it does not generate the data, which is the responsibility of each state government. Hence, there are differences in terms of information quality.

In particular, an important difference between the states is the information on the date each inmate enters and leaves the prison system. For this reason, we disregard these dates from the analysis and treat the data as a "registry of criminals" rather than a registry of inmates. In other words, we transform the panel data into a cross-section. We define a criminal as any individual who has committed a crime, was arrested, and, therefore, appears on our main dataset. Our focus is not on when the individual gets into the prison system but on whether he will be considered a criminal at some point in his life. DEPEN also provides information on their full names, their parent's full names, and inmates' date of birth, which — as we explain later — allow us to uniquely identify individuals and merge information from several administrative data.[6]

Initially, our data contains 3,256,204 criminals. However, we decide to follow the cohort born between 1996 and 2002. The idea is that, in this way, together with other datasets that we present later on, we can generate important descriptives — such as the age of entry in the prison system and the share of individuals that are persistent offenders — and also investigate school outcomes. This restriction leaves us with 330,612 remaining individuals.

As we explain in Section 2.2.2, we can identify all individuals born between 1996 and 2002 in Brazil based on the data provided by the Brazilian Tax Authority (*Receita Federal*). In this manner, we can compute, for each state, the share of criminals for the whole cohort of interest:

$$Share_s = \frac{Criminals_s}{All\_Individuals_s} \tag{2-1}$$

[5]The two states that we do not have information are Pará and Minas Gerais, which represent, respectively, around 4% and 10% of the Brazilian population. Hence, we still cover the states that encompass around 86% of Brazil's population.

[6]Although the father's name is commonly missing.

Where $Share_s$ is the share of criminals in state s for the 1996-2002 cohort; $Criminals_s$ is the total number of criminals born between 1996 and 2002 in state s; and $All\_Individuals_s$ is the total number of individuals born between 1996 and 2002 in state s. Figure 2.1 shows the geographical distribution of these shares. The mean share of criminals for the whole sample is 2.03% ($min = 0.69\%$ for Paraná; $max = 5.41\%$ for Acre). We also show the absolute number of criminals per state in Figure 2.2. The numbers above the bars are the share of criminals computed in equation 2-1.

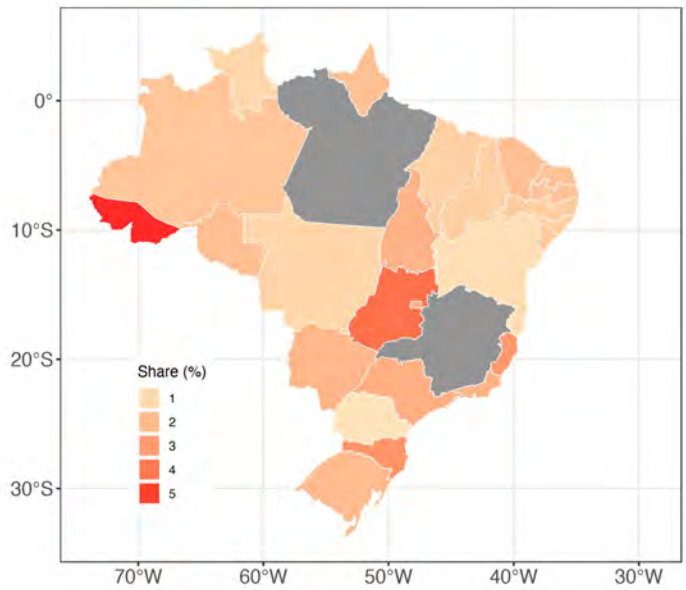

Figure 2.1: Share of Criminals Within The 1996-2002 Cohort
Note: The figure presents the geographical distribution of the share of criminals within the 1996-2002 cohort per state. This variable is constructed based on Equation 2-1, using all individuals born in this period.

2.2.2 Criminals Identification

To fully identify the criminals in our sample, we use data from the Brazilian Registry (*Cadastro de Pessoas Físicas*) provided by the Brazilian Tax Authority. It covers the universe of individuals in Brazil and contains information on their date of birth, gender, full name, their parent's full names, and their unique national identifier (CPF) — which allows us to merge inmates' information with other data without error.

To get the criminals' CPFs, we follow the same approach as Britto et al. (2022a). First, if the set of individual's name, mother's name, and date of birth is uniquely identified in the Brazilian population, we can identify the criminal's CPF. Next, we turn to other ways of uniquely identifying them. Specifically,

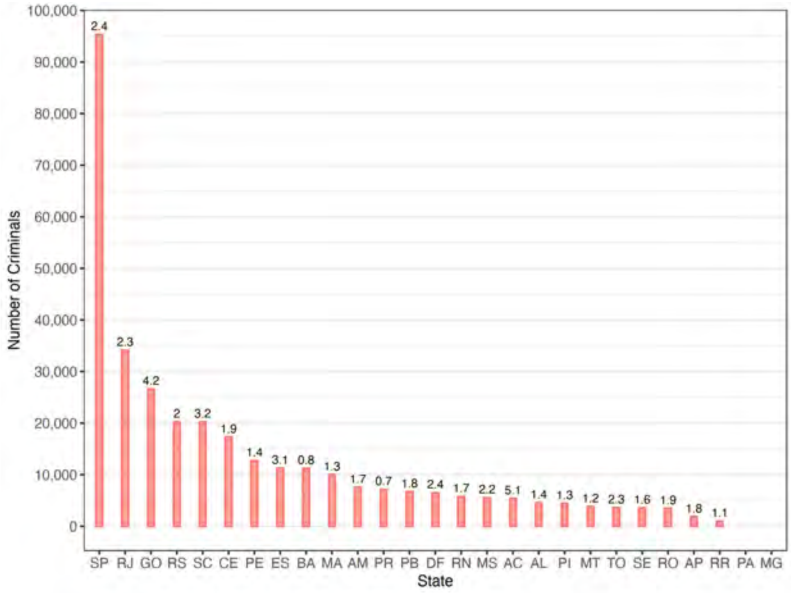

Figure 2.2: Number of Criminals Within The 1996-2002 Cohort

Note: The figure presents the total number of criminals within the 1996-2002 cohort per state. The number above each bar is the share of criminals within the 1996-2002 cohort per state, which is constructed based on Equation 2-1, using all individuals born in this period.

we join the datasets by the unique combination of name and date of birth, name and mother's name, and unique names.

This procedure allows us to retrieve the CPF of 79.5% of the individuals in our sample.[7] We drop the individuals we cannot identify, leaving us with a final dataset of 262,862 criminals. We follow a similar approach to recover the parents' CPFs of those inmates identified (55.2% for mothers and 35.4% for fathers).

Since our objective is to understand selection into criminality, we need a comparison group of non-criminals based on the general population. To construct this group, we draw a random sample from the Brazilian Registry and merge it with other administrative data.[8] The sample size ($n = 525,725$) equals twice the size of our final dataset on criminals to ensure statistical power. We weight the sample based on the population of each state. Hence, we have a final dataset of 788,587 individuals, encompassing criminals and non-criminals.

After retrieving their CPFs, we merge our data with a dataset — referred to as Administrative Records — that compiles demographic information

[7]4.2% of CPFs in the inmate's dataset were duplicated, and we randomly assigned one of these observations to remain in the sample.

[8]We first remove the individuals from the inmates' dataset to include only non-criminals in the random sample.

from formal employment records, the welfare registry, and election data. This dataset, constructed by Britto, Pinotti & Sampaio (2022), includes all individuals recorded in at least one of these sources. Specifically, they are included if they have at least one formal employment record between 2002 and 2019, are listed in the welfare registry from 2011 to 2020, or have candidacy records in any election from 2000 to 2020.

After successfully retrieving CPFs for 79.5% of the initial sample, we managed to merge approximately 64.77% (69.5% for the random sample) of these identified individuals with the Administrative Records. This results in about 51.5% (69.5%) of the original sample with demographic information, equivalent to 170,245 observations (365,285). Within this subsample, 92.2% of criminals are males (52.5%), and 61.66% are non-white (60.2%).

2.2.3 Other Data

After retrieving the national identifiers of the individuals in our dataset, we can merge our original data with several other datasets.

Demographic Census. To contrast the characteristics of the places where criminals come from compared to the general population, we collect socioeconomic data from the 2010 Brazilian demographic census. We retrieve the data on the level of the census tracts, the most granular level available, which is typically a block or a group of blocks. We have information such as the income per capita, the share of non-white people, whether it's an urban or rural neighborhood, and if the area is inside a *favela*.[9]

Thanks to the richness of our data, we can use the first zip code that appears in the Brazilian Registry as a proxy for individuals' childhood addresses and geolocate them using Google Maps API. Since we have the census tract shapefiles, we can perform a spatial join to identify which census tracts these zip codes belong to. Therefore, we have granular socioeconomic data on the level of the neighborhood in which criminals grow up.

Formal Employment. To construct measures of income and ties to the formal labor market of both criminals and their parents, we use *Relação Anual de Informações Sociais* (RAIS), an employer-employee dataset containing the universe of formal employments in Brazil for each year from 1985 to date. It contains information on each spell duration, the worker's wage, some demographics, information on the firm they work for, and their CPFs. RAIS is also one of the data sources used to construct the Administrative Records.

[9] *Favelas* are Brazilian slums typically characterized by high-density population, low-quality infrastructure and high-levels of poverty.

For the criminals' parents, we use this dataset for the period of 2002-2019 to gather information covering their children's ages from 0 to 18.[10] For the criminals, we focus on the period of 2014-2019 so we can identify their first tie with the formal labor market. 36% of criminals appear at least once in RAIS in this period, compared to 86% in the random sample.

School Census. We use the Brazilian School Census to explore criminals' school performance as part of our investigation of criminals' childhood and adolescence outcomes. The data comes from the Ministry of Education and covers the period from 2008 to 2017. It contains information on whether the individual was enrolled in a school, if they were in a lower grade than they should be if they were enrolled in a private school, and if they were in juvenile detention. We use the link between student identifiers and their CPFs established by Britto, Melo & Sampaio (2022), which is available only for students enrolled in 2014. These students may also be enrolled before and after this year; therefore, we use their last appearance in the school census to measure our variables of interest. The fact that we can only identify students enrolled in 2014 is one of the reasons why we restrict our sample to individuals born between 1996 and 2002.

National Standardized Exam. We also use another dataset from the Ministry of Education. The Saeb dataset has information on the 2011 national standardized exam on Language and Mathematics, which we use to correlate educational performance and criminal behavior. The exam is taken every two years in all public schools nationwide and on a sample basis in private schools. We link the student's CPFs through the student identifier used in the school census dataset.

Welfare Registry. To investigate the socioeconomic conditions of parents of criminals, we use the *Cadastro Único* (*CadÚnico*), which is the Federal Government welfare registry. Every social assistance beneficiary from 2011 to 2020 is included in this registry. The Welfare Registry is also one of the data sources used to construct the Administrative Records.

Criminal Charges. As discussed before, one of the limitations of our inmates' dataset is that we do not know when each individual enters and leaves the prison system. However, we have access to the universe of criminal charges presented in first-degree courts from 2009 to 2019, which we use to construct some important descriptives, although we do not use them in the main analysis.

[10]Because RAIS had a different layout before 2002, we start our analysis at this point, which leads some individuals to have this information starting when they are already 6 years old.

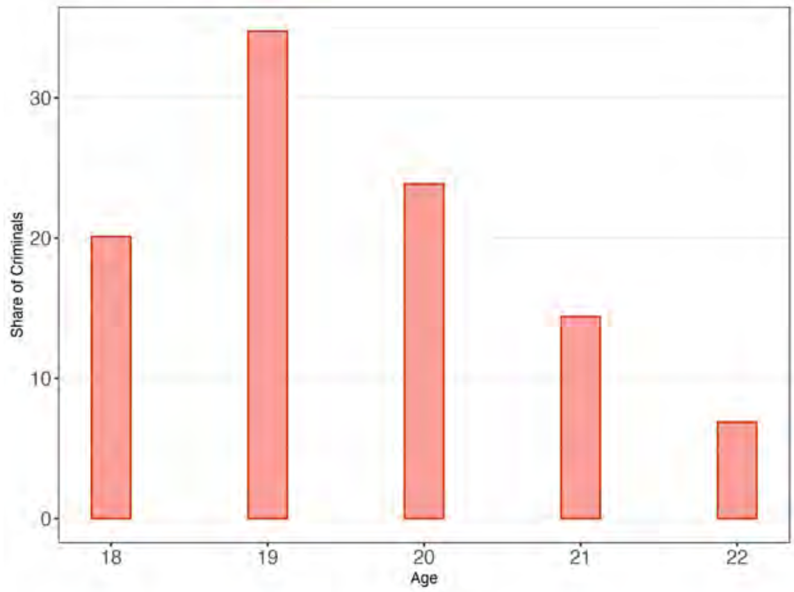

Figure 2.3: Age of First Criminal Charge Within The 1996-98 Cohort

Note: The figure presents the number of criminals that can be identified solely by their full name in Brazil and were born between 1996 and 1998, separating them by the age they had when they faced their first criminal charge.

Unlike the rest of our datasets, we do not have information on the individuals' CPFs, so we have to merge our data only with respect to the individuals who can be uniquely identified by their full names. Data on these criminal charges is public information. It was collected and made available by Kurier, a national firm providing paralegal services to law firms in Brazil.[11]

Using this restricted sample of individuals who are uniquely identified by their full names and also restricted to individuals who were born between 1996 and 1998, we can generate two important descriptives.[12] First, Figure 2.3 shows the age at which individuals first face criminal charges. We see that the peak of the first charge occurs when the individual is 19. In addition, we see in Figure 2.4 that over 56.7% of individuals face only one criminal charge in their early adulthood. Although these graphs are only proxies for criminal behavior, they are consistent with results from the literature, which finds that crime peaks before the individual reaches the age of 20 and that most criminals are not persistent offenders (CAF, 2014).

[11]This data was first used by Britto, Pinotti & Sampaio (2022).

[12]We restrict individuals born between 1996 and 1998 so our sample is not biased by individuals who could not have their first charges after 18 years old.

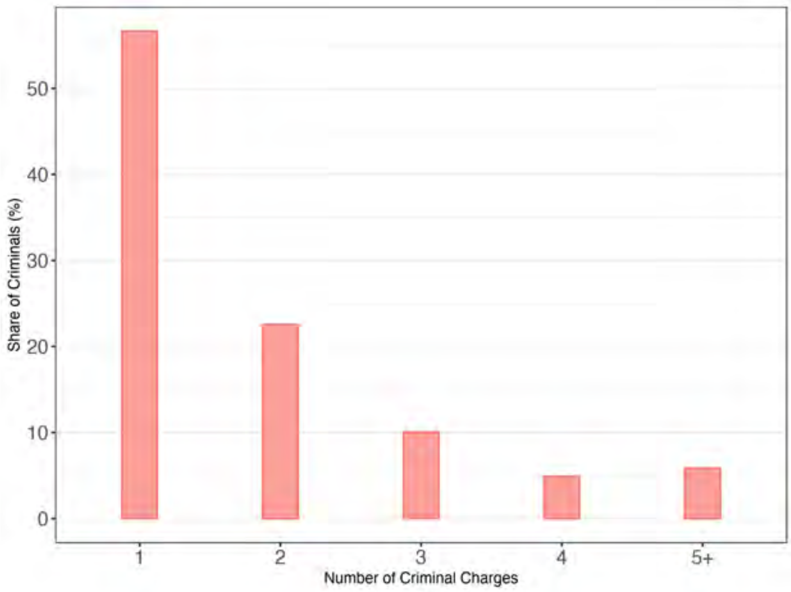

Figure 2.4: Number of Criminal Charges Within The 1996-98 Cohort

Note: The figure presents the number of criminals that can be identified solely by their full name in Brazil and were born between 1996 and 1998, separating them by the number of criminal charges they have faced.

3
Who Becomes a Criminal?

We now turn to the analysis of who becomes a criminal. Since our aim is to characterize the differences between criminals and non-criminals, we do not make any causal statements in all of the results we show in the paper. To proceed with the analysis, we run Equation 3-1:

$$\mathbf{I}(Criminal_i) = \alpha_0 + \boldsymbol{\beta}'\mathbf{X_i} + \lambda_1 \mathbf{I}(Male_i) + \lambda_2 \mathbf{I}(Non_White_i) + \phi_y + \gamma_n + \epsilon_i \quad (3\text{-}1)$$

Where $\mathbf{I}(Criminal_i)$ is a dummy indicating whether individual i is a criminal; $\mathbf{X_i}$ is a vector of covariates of interest for individual i; $\mathbf{I}(Male_i)$ and $\mathbf{I}(Race_i)$ are dummies indicating whether the individual is a male and if they are non-white; ϕ_y are year-of-birth fixed effects; and γ_n are neighborhood fixed effects.

Since we merge various datasets, many individuals lack some information from at least one of these datasets. For this reason, we present results for the sample in which all the information is available (Equation 3-1), and we also present results where we add dummies to indicate which observations are missing for each variable, and we input values for individuals who lack that information. We input zero for continuous variables, and for dummies, we randomly input 0 or 1 according to the frequency in the sample where the variable is not missing.

3.1
The Link Between Neighborhoods and Criminality

We start our analysis by investigating the neighborhoods where criminals have grown up. As discussed in Section 2.2.3, we gather information on the level of the census tract for the 2010 Brazilian Census. Hence, we measure neighborhood outcomes when individuals are between 8 and 14 years old. Since we do not have information on census tracts for all individuals in our full sample ($n = 788,587$), we present results for place characteristics for the sample where this information is available ($n = 158,111$). Table 3.1 presents the estimates for our covariates of neighborhood characteristics using state fixed effects instead of census tract fixed effects.[1]

[1]Including census tract fixed effects would eliminate the variation necessary to estimate the regressions.

Dependent Variable:	**I**(Criminal)		
Model:	(1)	(2)	(3)
Variables			
I(Urban)	0.152***	0.167***	0.103***
	(0.015)	(0.018)	(0.010)
I(Favela)	0.098***	0.078***	0.060***
	(0.006)	(0.004)	(0.004)
Share of Criminals	0.232***	0.243***	0.125***
	(0.008)	(0.007)	(0.007)
Log(Population)		0.010***	0.003
		(0.002)	(0.002)
Log(Per Capita Income)		-0.044***	0.016*
		(0.008)	(0.007)
Non-White			0.146***
			(0.016)
Missing Indicator	No	No	No
Mean Dep. Var.	0.349	0.349	0.349
Fixed-effects			
Year of Birth	Yes	Yes	Yes
State	Yes	Yes	Yes
Fit statistics			
Observations	158,111	158,111	158,111
Adjusted R^2	0.135	0.137	0.457

*Clustered (Year of Birth) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: The table presents information on the link between neighborhoods and criminality. The dependent variable equals one if the individual is a criminal and zero otherwise. We use the first address in the Brazilian Registry as a proxy for the places (census tracts) where individuals have grown up. We use data from the 2010 Brazilian Census for all variables except the share of criminals, which we construct based on our inmates dataset.

Table 3.1: Places

Due to the possibility of high collinearity between some of these variables, we present results in three steps (Columns 1-3). All the variables come from the Brazilian Census, except the share of criminals, which we construct based on our data on inmates. First, we are interested in the coefficients for urban areas, *favelas*, and the share of criminals in Column (1). All of them are statistically and economically significant. All else equal, growing up in an urban area is associated with a 15.2 percentage points (43.6%) higher chance of becoming a criminal. Similarly, growing up in a *favela* is associated with a 9.8 percentage points (28%) higher chance of becoming a criminal. These results suggest that being a criminal is an urban phenomenon, and these individuals are concentrated in more vulnerable places.

Also relevant is the coefficient for the share of criminals in the neighborhood. Comparing the place with the average share of criminals with the place with the lowest share of criminals, we see that, all else equal, an individual that grows up in the neighborhood with the average share has a 2.8 percentage points higher chance (or a 7.9% increase) in the probability of becoming a criminal.

In Column (2), we include the log of the population living in these neighborhoods and the log of their per capita income. First, we see that the coefficients for the other variables do not change. Next, the coefficients for both the population and the per capita income are both statistically but not economically significant. A 10% increase in both the population and the per capita income is associated to, respectively, an increase of 0.09 percentage points (0.27%) and a decrease of 0.42 percentage points (−1.2%) in the probability of criminal behavior. In addition, the adjusted R^2 virtually does not change with the inclusion of these variables (0.135 x 0.137). This possibly occurs because of collinearity between these variables, since urban places and *favelas* already capture variation related to population and income.

Finally, we include the share of non-white individuals as an additional covariate, which is a significant predictor. To illustrate its magnitude, if we compare someone who grows up in the place with the lowest share of non-white individuals with another individual who grows up in the place with the mean share of non-white individuals, the former would have a higher probability of becoming a criminal of 7.1 percentage points (20.4%). This also leads to the coefficient for the population to lose its statistical significance and the coefficient of per capita income to flip, although it continues without economic significance.

Our findings are consistent with results from the literature. They find that lower socioeconomic conditions may increase crime propensity by limiting the formation of emotional skills (CAF, 2014). Additionally, the correlation we find between being a criminal after 18 and the share of criminals in the neighborhoods where these individuals grow up has also been documented. Damm & Dustmann (2014), for instance, argue for the use of this variable instead of crime rates because the main mechanism between neighborhood crime and later criminal onset would be peer and role model effects, not exposure to crime itself.

A limitation of our results is that we consider the first zip code that appears in the Brazilian Registry. However, it may not reflect, in fact, the address where individuals were living between the ages of 8 and 14. To assess if this is a good proxy, it would be interesting to know the share of people who move at least once, given that they have lived in these neighborhoods. However, we do not have this statistic. Despite that, we argue that this is the best proxy one can get, given our data.

Overall, these results show that the places where criminals come from differ from the neighborhoods where non-criminals grow up. This fact suggests that we should indeed control for place fixed-effects when analyzing the other groups of variables.

3.2 Does Family Structure Matter in the Decision of Following a Criminal Path?

We now turn to what we call life's lottery. This group of variables consists of family characteristics that are virtually exogenous from the individual's point of view and may have an important impact on the opportunities they will have during life. Table 3.2 presents the results for these variables. Columns (1) and (2) do not include census tract fixed effects, while Columns (3) and (4) do. Columns (1) and (3) use the sample that has information on all these variables, while Columns (2) and (4) include dummies for observations with missing variables.

Results do not vary greatly based on the specification, and all variables are always statistically significant. The exception is the proxy of a present father in Column (3). This dummy is equal to one if we can identify the individuals' fathers. The coefficient for this variable varies from −0.3 percentage points (−1.2%) to −4.7 percentage points (−14%). In our preferred specification in Column (4), having a father present in one's life is associated with a 3.4 percentage points decrease (−9.6%) in the probability of becoming a criminal.

More relevant is the age at which the mother gives birth. This coefficient is statistically and economically significant in all specifications. Using our preferred specification, we have a decrease of 1.3 percentage points (−3.7%) for each additional year a mother waits to have the child. To illustrate, if a woman were to be a mother by the age of 20 instead of 16, this would be associated with a decrease in the probability of her child becoming a criminal by 5.2 percentage points (−14.7%). In addition, all else equal, an individual born in a family with 4 siblings *vis-à-vis* an individual with only 1 sibling has a 3.6 percentage points (10%) higher chance of becoming a criminal. Finally, although with a lower magnitude, the number of years of schooling completed by one's mother is also a relevant predictor of criminality in all specifications.

In our preferred model, an individual whose mother completed high school rather than only middle school has a 1.5 percentage points (-4.2%) lower chance of becoming a criminal.

Dependent Variable:	**I**(Criminal)			
Model:	(1)	(2)	(3)	(4)
Variables				
I(Father is Present)	-0.026***	-0.047***	-0.003	-0.034***
	(0.005)	(0.003)	(0.011)	(0.004)
Motherhood Age	-0.017***	-0.015***	-0.016*	-0.013***
	(0.002)	(0.001)	(0.006)	(0.001)
Motherhood Age Sqr.	0.000***	0.000***	0.000*	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Mother's Years of Schooling	-0.010***	-0.004***	-0.008***	-0.005***
	(0.001)	(0.000)	(0.002)	(0.000)
Number of Siblings	0.013***	0.014***	0.005*	0.012***
	(0.001)	(0.001)	(0.002)	(0.001)
Missing Indicator	No	Yes	No	Yes
Mean Dep. Var.	0.225	0.333	0.251	0.353
Fixed-effects				
Year of Birth	Yes	Yes	Yes	Yes
Census Tract			Yes	Yes
Fit statistics				
Observations	81,177	788,587	18,315	158,111
Adjusted R^2	0.470	0.318	0.564	0.439

*Clustered (Year of Birth) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: The table presents information on the family structure of individuals in our sample. The dependent variable equals one if the individual is a criminal and zero otherwise. Our proxy for present father is equal to one if the father can be identified and zero otherwise. Motherhood age is the age the mother was when the individuals in our sample were born. Mother's years of schooling come from the employer-employee dataset, and the number of siblings comes from the Brazilian Registry.

Table 3.2: Life's Lottery

In short, it seems that family structure is an important predictor of the decision to become a criminal. Families where both parents are present, there was no teenage pregnancy, the mother is educated, and there are not many mouths to feed are more likely not to have a member join a criminal path.

3.3 The Relationship Between Criminality and Family Income in Youth

Now, we continue to examine family outcomes during the individuals' childhoods, but we focus on their economic situation. Table 3.3 presents the results similarly to Table 3.2, but we separate the dummy variable indicating a formal job from the other variables, except when we impute values for missing data (Columns 3 and 6). The reason is that, in this case, the imputation can lead to collinearity since one needs to be in the formal labor market to have income from formal labor.

Although statistically significant, the economic significance of the coefficient on at least one parent having a formal job depends on the specification. Considering this variable alone, the coefficient varies from -0.8 percentage points (-2.4%) to -4.5 percentage points (-13.2%). On the other hand, conditional on the other regressors, it varies from -0.2 percentage points (-0.6%) to -0.8 percentage points (-2.3%), although this lowest estimate (in absolute value) does not consider place effects. Overall, we interpret these coefficients as weak evidence of a negative correlation between parents having a job during one's childhood and their later criminal onset.

More importantly, we see that the coefficient for parents who receive welfare when the individual is between 0 and 18 is significant in all our specifications, being a key predictor of criminal behavior. Specifically, it is associated with at least an increase in 5.7 percentage points (16.1%) in the probability of becoming a criminal. This is another strong evidence of a positive correlation between family economic vulnerability and the decision to become a criminal.

Next, we investigate whether having at least one parent with a stable job is associated with a reduction in the likelihood of a criminal trajectory. We do that by using public jobs as a proxy for job stability. This dummy takes the value of 1 if at least one of the parents works in the public sector at some point when the individual is between 0 and 10. Although the correlation sign is always negative for this variable, it is only statistically significant in Columns (2) and (3), losing its significance when we include neighborhood fixed effects.

We construct parental income for 4 age bins by summing the average formal income of both parents for each year within each bin. We also perform a Wald test for each specification to test if the labor income coefficients are jointly different from zero, which we cannot reject in all specifications. Although the labor income during the ages of 16-18 is the only age bin in which the coefficient is significant in all specifications, all age bins are relevant in our preferred model (Column 6). However, it does not seem that the parental

Dependent Variable:			**I**(Criminal)			
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
I(Parents Formal Job, Age 0-10)	-0.008*		-0.002**	-0.045***		-0.008*
	(0.003)		(0.001)	(0.006)		(0.003)
I(Parents Welfare, Age 0-18)		0.098***	0.070***		0.100**	0.057***
		(0.011)	(0.006)		(0.023)	(0.007)
I(Parents Public Job, Age 0-10)		-0.021*	-0.031***		-0.056	-0.008
		(0.010)	(0.004)		(0.032)	(0.016)
Log Labor Income (Age 0-5)		-0.017***	-0.009		-0.014**	-0.014**
		(0.002)	(0.005)		(0.003)	(0.003)
Log Labor Income (Age 6-10)		-0.015**	-0.018**		-0.009	-0.017**
		(0.003)	(0.005)		(0.005)	(0.006)
Log Labor Income (Age 11-15)		0.000	-0.011***		0.001	-0.010***
		(0.003)	(0.001)		(0.003)	(0.002)
Log Labor Income (Age 16-18)		-0.009***	-0.013***		-0.012***	-0.012***
		(0.002)	(0.001)		(0.002)	(0.001)
Missing Indicator	No	No	Yes	No	No	Yes
Mean Dep. Var.	0.333	0.270	0.333	0.342	0.301	0.353
Wald Test (L.I. Coefs.)		256	610		47	474
[p-value]		[0.000]	[0.000]		[0.000]	[0.000]
Fixed-effects						
Year of Birth	Yes	Yes	Yes	Yes	Yes	Yes
Census Tract				Yes	Yes	Yes
Fit statistics						
Observations	489,737	117,015	788,587	98,801	21,478	158,111
Adjusted R^2	0.051	0.455	0.317	0.245	0.550	0.438

*Clustered (Year of Birth) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: The table presents information on family income when the individuals in our sample were children or teenagers. The dependent variable equals one if the individual is a criminal and zero otherwise. The other variables come from the employer-employee dataset and the Welfare Registry. We compute Labor Income as the mean of the sum of the income by their mothers and fathers for each year within the child's age bin.

Table 3.3: Income

income from a formal job at different ages of childhood and adolescence is an important predictor. The magnitudes are rather small: if the parental formal income were to increase by 10% in all age bins, this would be equivalent to a decrease of only 0.5 percentage points (-1.4%) in the probability of becoming a criminal.

In sum, our results suggest that the most important predictor of criminal behavior related to the economic situation of parents is whether or not they receive welfare benefits, indicating a correlation between economic vulnerability and criminal behavior. On the other hand, parental income does not seem to be directly relevant. Nevertheless, our measure of income is restricted to formal jobs, which is an evident limitation since informality is a huge phenomenon in Brazil. It may be the case that if we used total income (formal plus informal income), we would have a more relevant coefficient, although we do not have this type of data available.

3.4 Youth Outcomes and the Prediction of Criminal Behavior

Our next step is to investigate the relationship between educational outcomes and future criminal behavior, which we do in Table 3.4. First, juvenile detention is a strong predictor of criminal behavior during adulthood, which is not surprising. This result suggests a persistence in criminal behavior from adolescence to adulthood. The coefficients for this variable are always significant and vary from 11 percentage points (31%), in our preferred specification in Column (6), to a striking 87 percentage points in Column (4). However, this highest coefficient should be interpreted carefully. One of the limitations of linear probability models is that they can lead to probabilities over 1, which is the case here. While this is not a feasible probability, it still stresses the strong relationship between juvenile detention and criminal behavior in adulthood.

Dependent Variable:			**I**(Criminal)			
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
I(Juvenile Detention)	0.368***	0.287***	0.186***	0.870***	0.252**	0.110***
	(0.043)	(0.014)	(0.023)	(0.056)	(0.081)	(0.013)
I(Educational Delay)	0.087***	0.144***	0.045***	0.070**	0.133***	0.042***
	(0.015)	(0.009)	(0.009)	(0.018)	(0.008)	(0.009)
I(Private School)	-0.018	-0.032***	-0.068***	-0.062	-0.002	-0.064***
	(0.013)	(0.003)	(0.009)	(0.071)	(0.054)	(0.006)
Std. Score 5th Gr.	-0.021***		-0.016	-0.014		-0.017**
	(0.003)		(0.010)	(0.010)		(0.006)
Std. Score 9th Gr.		-0.041***	-0.055***		-0.035***	-0.056***
		(0.002)	(0.003)		(0.004)	(0.004)
Missing Indicator	No	No	Yes	No	No	Yes
Mean Dep. Var.	0.204	0.202	0.333	0.203	0.232	0.353
Fixed-effects						
Year of Birth	Yes	Yes	Yes	Yes	Yes	Yes
Census Tract				Yes	Yes	Yes
Fit statistics						
Observations	35,737	30,089	788,587	5,674	7,311	158,111
Adjusted R^2	0.413	0.513	0.330	0.501	0.622	0.446

*Clustered (Year of Birth) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: The table presents information on outcomes during adolescence, focusing on educational outcomes. The dependent variable equals one if the individual is a criminal and zero otherwise. Data for the other variables comes from the Brazilian School Census and the 2011 national standardized exam. Information is restricted to individuals who were enrolled in a school in 2014. Education delay equals one if the individual is in a grade below the expected based on their age. Standard Scores were constructed by summing their scores in Maths and Language and then standardizing it to a N(0,1).

Table 3.4: School Outcomes

Another important predictor is whether or not the individual was in a grade below the one they should have been based on their age. The coefficient for educational delay is always significant and varies from 4.2 percentage points (11.9%) in our preferred specification to 14.4 percentage points (71%) in a specification without neighborhood fixed effects. We interpret this result as an evidence that criminals perform worse in school.

This is also reinforced when we observe the coefficients for the 2011 national standardized exam scores. We sum the scores on Language and Maths separately for the 5th and 9th grades, and then standardize the resulting scores. As individuals only have a score for one of each grade, we separate the regressions for each grade, except in Columns (3) and (6), where we include the missing variable dummies. We find that the score in the 9th grade is a more important predictor than the score for the 5th grade. It has a higher magnitude and is always significant, while in some specifications, the coefficient for the 5th grade is not. In particular, in our preferred specification, a 1 standard-deviation increase in the score of the exam for the 9th grade is associated with a decrease of 5.6 percentage points (-15.9%) in the chance of becoming a criminal, while the equivalent for the 5th grade is 1.7 percentage points (-4.8%).

Finally, we also explore the relationship between being a student in a private school and criminality. In Brazil, private schools are, on average, better than public schools. In addition, their students are typically richer and have better opportunities than their peers in public schools. We find that, indeed, there seems to exist a negative relationship between criminality and studying in a private school. However, this result is not significant in all specifications, and in the case of Column (5), the magnitude of the coefficient is not relevant. Despite that, in our preferred model, we find that studying in a private school is associated with a reduction in the probability of becoming a criminal by 6.4 percentage points (-18.1%), which is a considerable correlation.

Overall, we find that criminal behavior in early adulthood is correlated with worse educational outcomes and that individuals with previous contact with the juvenile detention system are more likely to become criminals later on. Both of these results are consistent with previous findings in the literature. Draca & Machin (2015) and CAF (2014), for instance, argue that education can reduce criminal behavior by increasing earnings in the legal market and by developing cognitive and non-cognitive skills that reduce crime propensity. In relation to the finding of juvenile detention, Sampson & Laub (2003) also discuss that early criminal activity is associated to a longer criminal career.

3.5 Adult Outcomes and Their Relationship to Criminality

We end our analysis by investigating outcomes during adulthood in Table 3.5. The literature highlights the importance of adult outcomes in predicting criminal behavior compared to only analyzing outcomes during childhood (Laub & Sampson, 1993; Sampson & Laub, 2005). Most of our focus here is on labor market outcomes. A limitation of our data is that we do not know the exact timing of the individual's arrest; we only know that they were arrested at some point between 18 and their age in 2019. In this manner, our labor market measures do not take into account whether the individual had been arrested before or after they were measured. Despite that, the analysis can still point to some interesting selection patterns, and we proceed with it.

Dependent Variable:			**I**(Criminal)			
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
I(Formal Job)	-0.036**		0.165***	-0.079***		0.110**
	(0.011)		(0.036)	(0.007)		(0.036)
School Years	-0.039***	-0.020***	-0.022***	-0.034***	-0.019***	-0.020***
	(0.005)	(0.002)	(0.002)	(0.004)	(0.002)	(0.001)
Log(Labor Income)		-0.026***	-0.030***		-0.023***	-0.028***
		(0.005)	(0.004)		(0.005)	(0.005)
Spell Duration		-0.002***	-0.002***		-0.001***	-0.001***
		(0.000)	(0.000)		(0.000)	(0.000)
I(Public Job)		-0.096***	-0.051**		-0.075***	-0.034***
		(0.011)	(0.014)		(0.011)	(0.008)
Missing Indicator	No	No	Yes	No	No	Yes
Mean Dep. Var.	0.333	0.303	0.333	0.346	0.341	0.353
Fixed-effects						
Year of Birth	Yes	Yes	Yes	Yes	Yes	Yes
Census Tract				Yes	Yes	Yes
Fit statistics						
Observations	535,530	289,599	788,587	109,669	64,314	158,111
Adjusted R^2	0.110	0.496	0.350	0.286	0.555	0.465

*Clustered (Year of Birth) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: The table presents information on adult outcomes, focusing on labor market outcomes. The dependent variable equals one if the individual is a criminal and zero otherwise. The other variables come from the employer-employee dataset.

Table 3.5: Labor Market Outcomes

We start by analyzing the correlation between having a formal job and becoming a criminal. We create a dummy to indicate whether an individual has at least one formal job between 18 and their age in 2019. When we only include this variable and years of schooling in the regression (Columns 1 and 4), there appears to be a considerable negative correlation between having a formal job and criminal activity. In both specifications, having a formal job reduces the probability of criminal behavior in, respectively, 3.6 percentage points (−10.8%) and 7.9 percentage points (−22.8%). Again, we stress that this is only a correlation, it can be the case that having a job indeed reduces criminal behavior; that criminal behavior reduces formal employment; or it could go both ways.

Nevertheless, when we include other labor market outcomes in the regression, the sign of the coefficients for formal employment flips. However, this is likely due to collinearity between this variable and the other labor market outcomes due to the imputation process.[2] Hence, we argue that the estimates in Columns (1) and (4) are more likely to reflect the true correlation between the dependent variable and having a formal job. This is reinforced by the R^2 in Column (4), which suggests that this variable has a high prediction power over the decision to become a criminal.

We also see that the higher the number of school years, the lower the chance of engaging in criminal behavior. In particular, in Column (5) — our preferred specification due to the problem of collinearity we mentioned before —, each additional year of schooling is equivalent to a reduction of 1.9 percentage points (−5.6%) in the probability of criminal activity. To illustrate, a person who completed high school compared to a person who only completed middle school has a 5.7 p.p (−16.7%) lower chance of becoming a criminal, all else equal. Similarly to what we discussed in the previous section, this result of schooling is consistent with Lochner & Moretti (2004), which show that additional years of education can reduce criminal behavior.

In relation to the formal income, the coefficient is also rather stable but not very relevant in terms of its magnitude. In Column (5), a 10% increase in one's wage is equivalent to a reduction of only 0.22 percentage points (−0.64%) in the probability of criminal activity. In addition, the spell duration — which we use as a proxy for volatility in the formal market — is not very economically significant. However, the sign of the variable goes in the expected direction in all specifications. Finally, as a measure of job stability, we include a dummy in the regression indicating if the individual has at least one public job within this period. The estimate for this variable is always significant both statistically and economically. In particular, in our preferred specification, having a public job is associated with a considerable 7.5 percentage points decrease (−21.9%)

[2]Individuals that the dummy for formal employment equals zero will have zero inputted for the other variables, while individuals that have at least one formal employment in the period will have the information for all other variables. Since the proportion of criminals that this dummy equals zero is much higher than the share of non-criminals, the coefficient for formal employment is expected to flip.

in the chance of criminal behavior.

Taken together, our results seem to be aligned with the previous literature. Laub & Sampson (1993), for instance, stress the importance of social bonds in adulthood such as having a job. The idea is that having a job increases social bonds and also restricts the time available for committing a crime, acting as an incapacitation effect (Laub & Sampson, 1993; Nguyen & Loughran, 2018). Also relevant is the potential "scarring effects" that crime can have on future labor market opportunities, which can lead to an increased chance of recidivism (Draca & Machin, 2015). On the other hand, the finding on labor income, although in the same direction, is quite weak and different from other results in the literature (Draca & Machin, 2015).

4
Conclusion

We document the profile of criminals, contrasting them with the general population. We investigate the characteristics of the places where these individuals grow up, their family configurations, family economic situations during their childhood and adolescence, and their educational and labor market outcomes. Our work is descriptive, and our findings are consistent with previous results from the literature. We expand on this by creating a comprehensive profile of criminals, considering a broad and integrated set of characteristics for the same individuals without focusing on specific channels. Additionally, we contribute to the literature by providing evidence from a developing country, an area that has been underexplored.

Our findings suggest that criminals come from more urban and vulnerable places and are more exposed to other criminals. In addition, our results indicate that family structure matters and that parents of criminals are more economically vulnerable than parents of non-criminals. Finally, criminals have worse educational performance, a more often history of juvenile detention, and worse labor market performance.

In sum, we find that poor socioeconomic conditions during childhood and adolescence are highly correlated with criminality and poor socioeconomic conditions in early adulthood. Our analysis shows that both Situational Action Theory (Wikström, 2009) and Age-Graded Theory (Laub & Sampson, 1993; Sampson & Laub, 2003; Sampson & Laub, 2005) can explain part of the story of following a criminal path. However, we stress that these theories should be interpreted in a complementary manner.

Extensions of this work should investigate if there is heterogeneity in relation to the type of crimes committed. It may be the case that the profile of individuals arrested by murder is considerably different from that of robbers. In addition, it would be interesting to contrast "common" criminals with gang members to understand if there is any difference in the determinants of selection into gang membership *vis-à-vis* selection into "common" criminality.

5
Bibliography

Aizer, A. Neighborhood violence and urban youth. **NBER Working Paper**, 2008. Cited in page 13.

Anderson, D. A. The aggregate burden of crime. **Journal of Law and Economics**, University of Chicago Press, v. 42, p. 611–642, 1999. ISSN 00222186. Cited in page 11.

Barr, A. & Gibbs, C. R. Breaking the cycle? intergenerational effects of an anti-poverty program in early childhood *. **Accepted, Journal of Political Economy**, 2022. Cited in page 13.

Becker, G. S. Crime and punishment: An economic approach. **Journal of Political Economy**, v. 76, p. 169–217, 1968. Cited 2 times in pages 11 and 13.

Bell, A. et al. Who becomes an inventor in america? the importance of exposure to innovation. **The Quarterly Journal of Economics**, 2018. Disponível em: <https://academic.oup.com/qje/advance-article-abstract/doi/10.1093/qje/qjy028/5218522>. Cited in page 14.

Britto, D. G. C. et al. Intergenerational mobility in the land of inequality *. **Working Paper**, 2022. Cited in page 18.

Britto, D. G. C. et al. Parenthood, crime and domestic violence in brazil. **Working Paper**, 2022. Cited in page 13.

Britto, D. G. C. & Melo, C. & Sampaio, B. The kids aren't alright: Parental job loss and children's outcomes within and beyond schools *. **Working Paper**, 2022. Cited 2 times in pages 13 and 21.

Britto, D. G. C. & Pinotti, P. & Sampaio, B. The effect of job loss and unemployment insurance on crime in brazil. **Econometrica**, v. 90, p. 1393–1423, 2022. Cited 3 times in pages 13, 20, and 22.

Bó, E. D. et al. Who becomes a politician? **The Quarterly Journal of Economics**, v. 132, p. 1877–1914, 6 2017. Cited in page 14.

CAF. Towards a safer latin america a new perspective to prevent and control crime. **Development Bank of Latin America**, 2014. Cited 4 times in pages 14, 22, 26, and 32.

Campaniello, N. & Gray, R. & Mastrobuoni, G. Returns to education in criminal organizations: Did going to college help michael corleone? **Economics of Education Review**, Elsevier Ltd, v. 54, p. 242–258, 10 2016. ISSN 02727757. Cited in page 13.

Carvalho, L. S. & Soares, R. R. Living on the edge: Youth entry, career and exit in drug-selling gangs. **Journal of Economic Behavior and Organization**, Elsevier, v. 121, p. 77–98, 1 2016. ISSN 01672681. Cited in page 13.

Chetty, R. & Hendren, N. The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects *. **The Quarterly Journal of Economics**, v. 133, p. 1163–1228, 2018. Cited in page 13.

Damm, A. P. & Dustmann, C. Does growing up in a high crime neighborhood affect youth criminal behavior? **American Economic Review**, American Economic Association, v. 104, p. 1806–1832, 2014. ISSN 00028282. Cited 3 times in pages 11, 13, and 26.

Draca, M. & Machin, S. Crime and economic incentives. **Annual Review of Economics**, Annual Reviews Inc., v. 7, p. 389–408, 8 2015. ISSN 19411391. Cited 3 times in pages 13, 32, and 35.

Ehrlich, I. Participation in illegitimate activities: A theoretical and empirical investigation. **Journal of Political Economy**, v. 81, p. 521–565, 1973. Cited in page 13.

Fair, H. & Walmsley, R. World prison population list 13th edition. **Institute for Crime Justice Policy Research**, 2021. Cited in page 16.

Fella, G. & Gallipoli, G. Education and crime over the life cycle. **Review of Economic Studies**, Oxford University Press, v. 81, p. 1484–1517, 10 2014. ISSN 1467937X. Cited in page 13.

Heckman, J. et al. Analyzing social experiments as implemented: A reexamination of the evidence from the highscope perry preschool program. **Quantitative Economics**, The Econometric Society, v. 1, p. 1–46, 2010. ISSN 1759-7323. Cited in page 13.

INFOPEN. **Levantamento Nacional de Informações Penitenciárias**. 2014. Cited in page 16.

Koskela, E. & Virén, M. An occupational choice model of crime switching. **Applied Economics**, Routledge, v. 29, p. 655–660, 1997. ISSN 00036846. Cited in page 13.

Laub, J. H. & Sampson, R. J. Turning points in the life course: Why change matters to the study of crime. **Criminology**, v. 31, p. 301–325, 1993. ISSN 17459125. Cited 5 times in pages 11, 14, 33, 35, and 36.

Levine, R. & Rubinstein, Y. Smart and illicit: Who becomes an entrepreneur and do they earn more? **The Quarterly Journal of Economics**, 2016. Disponível em: <http://qje.oxfordjournals.org/>. Cited in page 14.

Levitt, S. D. & Venkatesh, S. A. An economic analysis of a drug-selling gang's finance. **The Quarterly Journal of Economics**, v. 115, p. 755–789, 2000. Disponível em: <http://qje.oxfordjournals.org/>. Cited in page 13.

Levitt, S. D. & Venkatesh, S. A. Growing up in the projects: The economic lives of a cohort of men who came of age in chicago public housing. **American Economic Review**, v. 91, p. 79–84, 2001. Cited 2 times in pages 11 and 13.

Lochner, L. **Education, work, and crime: A human capital approach**. 2004. 811-843 p. Cited in page 13.

Lochner, L. & Moretti, E. The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. **American Economic Review**, v. 94, p. 155–189, 2004. Disponível em: <http://www.econ.ucla.edu/moretti/papers.html.> Cited 2 times in pages 13 and 34.

Mata, A. D. & Dulmen, M. H. van. Adult-onset antisocial behavior trajectories: Associations with adolescent family processes and emerging adulthood functioning. **Journal of Interpersonal Violence**, v. 27, p. 177–193, 1 2012. ISSN 08862605. Cited in page 14.

Monteiro, J. et al. Criminal enterprises: Evidence from rio de janeiro. **Working Paper**, 2022. Cited in page 16.

Monteiro, J. & Rocha, R. Drug battles and school achievement: Evidence from rio de janeiro's favelas. **Review of Economics and Statistics**, MIT Press Journals, v. 99, p. 213–228, 5 2017. ISSN 15309142. Cited in page 11.

Murray, J. & Farrington, D. P. Parental imprisonment: effects on boys' antisocial behaviour and delinquency through the life-course. **Journal of child psychology and psychiatry, and allied disciplines**, v. 46, p. 1269–1278, 2005. ISSN 00219630. Cited in page 14.

Nguyen, H. & Loughran, T. A. On the measurement and identification of turning points in criminology. **Annual Review of Criminology**, p. 335–358, 2018. Disponível em: <https://doi.org/10.1146/annurev-criminol->. Cited 2 times in pages 14 and 35.

Sampson, R. J. & Laub, J. H. Desistance from crime over the life course*. **In Handbook of the life course**, Boston, MA. Springer, p. 295–309, 2003. Cited 3 times in pages 14, 32, and 36.

Sampson, R. J. & Laub, J. H. A life-course view of the development of crime. In: . [S.l.: s.n.], 2005. v. 602, p. 12–45. ISSN 00027162. Cited 4 times in pages 11, 14, 33, and 36.

Sharkey, P. The long reach of violence: A broader perspective on data, theory, and evidence on the prevalence and consequences of exposure to violence. **Annual Review of Criminology**, v. 1, p. 85–102, 2017. Disponível em: <https://doi.org/10.1146/annurev-criminol->. Cited in page 13.

Sviatschi, M. M. Making a narco: Childhood exposure to illegal labor markets and criminal life paths. **Econometrica**, v. 90, p. 1835–1878, 7 2022. Disponível em: <www.micaelasviatschi.com/research/.> Cited 2 times in pages 11 and 13.

UNODC. Global study on homicide: Homicide trends, patterns and criminal justice response. **United Nations Office on Drugs and Crime**, Booklet 2, 2019. Cited in page 16.

Wikström, P.-O. H. Crime propensity, criminogenic exposure and crime involvement in early to mid adolescence. **Monatsschrift für Kriminologie und Strafrechtsreform**, v. 92, p. 253–266, 2009. Disponível em: <www.pads.ac.uk>. Cited 3 times in pages 11, 14, and 36.

Zara, G. & Farrington, D. P. Childhood and adolescent predictors of late onset criminal careers. **Journal of Youth and Adolescence**, v. 38, p. 287–300, 3 2009. ISSN 00472891. Cited in page 14.