UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA
DOUTORADO EM ECONOMIA

LUCICLEYTON HENRIQUE DE FARIAS

ESSAYS ON APPLIED ECONOMETRICS

RECIFE – PERNAMBUCO
2023

LUCICLEYTON HENRIQUE DE FARIAS

ESSAYS ON APPLIED ECONOMETRICS

Tese apresentada ao Curso de Doutorado em Economia do Programa de Pós-Graduação em Economia do Centro de Ciências Sociais Aplicadas da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de doutor em Economia. Área de Concentração: Economia

Orientador: Prof. Dr. Breno Ramos Sampaio

RECIFE – PERNAMBUCO

2023

Catalogação na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

F224e Farias, Lucicleyton Henrique de
Essays on applied econometrics / Lucicleyton Henrique de Farias. - 2023.
88 folhas: il. 30 cm.

Orientador: Prof. Dr. Breno Ramos Sampaio.
Tese (Doutorado em Economia) – Universidade Federal de Pernambuco, CCSA, 2023.
Inclui referências.

1. Econometria. 2. Crime. 3. Brasil. I. Sampaio, Breno Ramos (Orientador). II. Título.

336 CDD (22. ed.) UFPE (CSA 2023 – 085)

LUCICLEYTON HENRIQUE DE FARIAS

ESSAYS ON APPLIED ECONOMETRICS

Tese apresentada ao Curso de Doutorado em Economia do Programa de Pós-Graduação em Economia do Centro de Ciências Sociais Aplicadas da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de doutor em Economia. Área de Concentração: Economia

Aprovada em: 14 de Julho de 2023

BANCA EXAMINADORA

Prof. Dr. Breno Ramos Sampaio (Orientador)
Universidade de Federal de Pernambuco - UFPE

Prof. Dr. Gustavo Ramos Sampaio (Examinador Interno)
Universidade Federal de Pernambuco - UFPE

Prof. Dr. Giuseppe Trevisan Cruz (Examinador Interno)
Universidade Federal de Pernambuco - UFPE

Prof. Dr. Vítor Possebom (Examinador Externo)
Fundação Getúlio Vargas - FGV/EESP

Prof. Dr. Antonio Vinicius Barros Barbosa (Examinador Externo)
Universidade Federal da Paraíba - UFPB

Este trabalho é dedicado a todas as pessoas que têm sido parte fundamental em minha jornada até este momento: minha família, professores e amigos. Que este trabalho possa contribuir para o avanço do conhecimento e o debate público.

AGRADECIMENTOS

Nesta jornada de pesquisa e aprendizado, fui abençoado por contar com o apoio e a influência positiva de muitas pessoas extraordinárias. Gostaria de expressar minha gratidão a todos aqueles que contribuíram para a realização desta tese.

Primeiramente, quero expressar minha profunda gratidão ao meu orientador, Prof. Dr. Breno Ramos Sampaio, por sua orientação habilidosa, sabedoria e paciência incansável ao longo de todo o processo de pesquisa. Sua dedicação à minha formação acadêmica foi inestimável. Mesmo nos momentos mais desafiadores, pude contar com seu apoio. Suas críticas construtivas moldaram este trabalho e minha carreira profissional de maneiras que ultrapassam as palavras.

Minha família merece um agradecimento especial. Ao meu Pai, Josenildo, e em especial minha Mãe, Lucicleide, que sempre me apoiaram incondicionalmente. Seu amor e incentivo foram minha força motriz. O exemplo de determinação e os valores que eles me transmitiram moldaram a pessoa que me tornei.

À minha esposa, Maria Carolina, que esteve comigo desde o início com suporte incondicional e acreditando em mim mais do que eu acreditava em mim mesmo. Sua paciência, amor e compreensão foram minha âncora nos momentos de incerteza. Seu apoio me inspirou a continuar buscando meus objetivos, mesmo quando os desafios pareciam insuperáveis.

Aos meus amigos, colegas de estudo, por compartilharem suas ideias, discussões enriquecedoras e pelo apoio mútuo durante a difícil trajetória acadêmica. Nossas trocas foram um componente vital para minha formação acadêmica e pessoal.

Às pessoas que generosamente dedicaram seu tempo para participar de discussões no Grupo de Avaliação de Políticas Públicas e Econômicas (GAPPE). Nossas conversas foram cruciais para a obtenção dos dados e insights que sustentam este trabalho.

Por fim, gostaria de agradecer à Universidade Federal de Pernambuco e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo financiamento e pela oferta de recursos que viabilizaram esta pesquisa.

Este trabalho é o resultado de esforços coletivos, e estou honrado por ter sido rodeado por indivíduos tão notáveis. Que este estudo contribua de alguma forma para o avanço do conhecimento e para o debate público em questão.

Com sincera gratidão, Lucicleyton Henrique de Farias.

“It always seems impossible until it’s done.”

(Nelson Mandella)

RESUMO

Essa tese contribui para o campo recente que estuda a interseção entre o sistema judicial e os resultados socioeconômicos individuais. Utilizando modelos econométricos e um conjunto de dados único composto por casos criminais e fontes de dados administrativos no Brasil, esta pesquisa tem como objetivo abordar duas questões cruciais que têm importantes implicações políticas.

Capítulo 1. Conviction, Employment, and Recidivism: Evidence from Brazil

Este artigo examina o impacto de uma condenação criminal nos resultados no mercado de trabalho e reincidência no Brasil, usando uma abordagem de variável instrumental. Nossos resultados mostram que uma condenação criminal reduz significativamente o emprego em 22% e os ganhos salariais em 25% dentro de três anos após o início do processo. Também encontramos evidências robustas de que uma condenação criminal aumenta a atividade criminal subsequente em 13 pontos percentuais. Nossa análise de heterogeneidade mostra que esses efeitos estão concentrados entre indivíduos acusados de crimes de baixa gravidade. Esses resultados sugerem que o estigma social pode desempenhar um papel significativo no mercado de trabalho. Nosso estudo fornece as primeiras evidências causais dos efeitos diretos de uma condenação criminal sobre empregabilidade e reincidência em um contexto de um país não desenvolvido.

Palavras-chaves: Variável Instrumental; Randomização; Crime; Brasil; Condenação;

Capítulo 2. Conviction's Echo: Unveiling Family Consequences in Brazil

Este artigo utiliza uma abordagem de variável instrumental para investigar o impacto da condenação de um indivíduo sobre os resultados no mercado de trabalho, reincidência criminal e educação dos membros da família no Brasil. Nossos resultados apontam reduções significativas no desempenho do mercado de trabalho para os membros da família afetadas em termos de empregabilidade e de ganhos salariais. Além disso, nossa análise demonstra que a condenação aumenta a atividade criminal, especialmente em relação a crimes graves. Embora tenham sido estimados efeitos adversos sobre educação, os resultados não apresentaram significância estatística. Além disso, a análise de heterogeneidade mostra que os efeitos adversos estão concentrados entre os homens e irmãos. Este estudo fornece as primeiras evidências causais das consequências de uma condenação criminal sobre os membros da família no Brasil.

Palavras-Chaves: Variável Instrumental; Randomização; Crime; Brasil; Condenação; Família;

ABSTRACT

This thesis contributes to the recent and evolving field that studies the intersection between the judicial system and individuals' socioeconomic outcomes. Leveraging the power of econometric models and using a unique dataset comprising criminal cases and extensive administrative data sources from Brazil, this research aims to address two crucial questions that hold significant policy implications.

Chapter 1. Conviction, Employment, and Recidivism: Evidence from Brazil

This paper examines the impact of a criminal conviction on labor market outcomes and recidivism in Brazil, using an instrumental variable approach. Our findings show a criminal conviction significantly reduces employment by 22 percent, and earnings by 25 percent within three years after the case starts. We also find evidence that a criminal conviction increases following criminal activity by 13 percentage points. Our heterogeneity analysis shows that these adverse effects are concentrated among individuals charged with low-severity crimes. These results suggest that social stigma might play a significant role in the negative consequences of criminal records on labor market prospects. Our study provides the first causal evidence of the direct effects of a criminal conviction on labor and recidivism outcomes in a non-developed country context.

Keywords: Instrumental Variable; Random Assignment; Crime; Brazil; Conviction;

Chapter 2. Conviction's Echo: Unveiling Family Consequences in Brazil

This paper employs an instrumental variable approach to investigate the impact of an individual's conviction on labor market outcomes, criminal behavior, and education among family members in Brazil. Our findings reveal compelling evidence of significant reductions in labor market performance for affected family members across both extensive (employment rates) and intensive (earnings) margins. Additionally, our analysis demonstrates conviction increases subsequent criminal activity, particularly in relation to serious offenses. While adverse effects on educational outcomes were estimated, the results did not yield statistical significance. Furthermore, heterogeneity analysis shows adverse effects are concentrated among males and siblings. This study provides the first causal evidence of the consequences stemming from individual's conviction on family members in Brazil. By establishing a causal link, it offers compelling evidence of the far-reaching impact of such convictions, shedding light on the pervasive nature of the consequences involved.

Keywords: Instrumental Variable; Random Assignment; Crime; Brazil; Conviction; Family;

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 CONVICTION, EMPLOYMENT AND RECIDIVISM: EVIDENCE FROM BRAZIL

1.1 INTRODUCTION

The criminal justice system plays a crucial role in shaping the lives of individuals who have been convicted of a crime. While the immediate consequences of legal sanctions, such as imprisonment or fines, are well-documented, the long-term effects of a criminal record on an individual's life remain less understood. In recent years, there has been a notable global increase in the incarcerated population[1]. Applied econometrics has seen a growing body of research dedicated to examining this topic. However, the findings from these studies have yielded mixed results. Some studies indicate no significant impact (KLING, 2006; GREEN; WINIK, 2010; LOEFFLER, 2013; DOBBIE *et al.*, 2019; GARIN *et al.*, 2022), while others have found negative consequences (AIZER; JR, 2015; MUELLER-SMITH, 2015; HARDING *et al.*, 2018), and yet another set of studies have uncovered positive outcomes (BHULLER *et al.*, 2020; ARTEAGA, 2021; NORRIS; PECENCO; WEAVER, 2021). Despite the high incarceration rate in Latin America, there remains a significant gap in understanding the direct effects on convicts in the region. It is, therefore, crucial to address this gap and gain a more comprehensive understanding of the consequences of criminal records in a context where crime rates surpass those of developed countries and the State's control over its territory may be limited.

Estimating the causal impact of a criminal conviction poses several challenges. First, the lack of individual-level panel data on criminal records makes it difficult to track individuals' criminal histories over time. Without such data, it becomes challenging to investigate the consequences of a conviction. Additionally, identifying a causal relationship between a criminal conviction and labor and subsequent criminal outcomes is complicated by endogeneity issues. For instance, criminal convicted individuals may be systematically different from non-convicted ones in terms of unobservable characteristics. As a result, identifying the true effect of a criminal record on subsequent outcomes requires rigorous research designs that can account for these endogeneity problems.

In this paper, we investigate the causal effect of a criminal conviction on labor and recidivism outcomes in Brazil. We address the data challenge by linking a rich set of collected criminal cases in Brazil to the universe of formal workers and firms (*RAIS*) to construct a unique

[1] According to the World Prison Population List, available at www.prisonstudies.org, the number of individuals in prison has grown by 24% from 2000-2021, with the most significant increases seen in South America (200%) and Southeast Asia (116%).

panel dataset that allows us to both track labor and recidivism outcomes for individuals charged with criminal offenses. Also, we overcome the endogeneity issue by exploiting the institutional rule that dictates the random assignment of judicial cases to courtrooms that differ systematically in their tendency to convict. We construct the courtroom stringency measure as the leave-one-out average of the conviction rate and use it as an instrument for conviction decisions. By applying this instrumental variable design, we can estimate the local average treatment effect of a criminal conviction on labor and recidivism outcomes in Brazil.

Our study offers important findings on the impacts of criminal convictions. Through the use of an instrumental variable design, we provide strong evidence that conviction reduces the likelihood of ever working by 22 percent, total days worked by 64 percent, and total earnings by 26 percent within three years after the case starts. However, the underlying mechanisms that drive these adverse effects remain unclear. One potential explanation is that social stigma linked to having a criminal record may impede these individuals from obtaining employment opportunities. On the other hand, the incapacitation of convicted individuals, who are more likely to be incarcerated, may limit their employment prospects.

Determining the contribution of incapacitation to the adverse effects of criminal convictions is complicated due to the lack of data on the prison time of a specific conviction. However, If incapacitation plays a considerable role, convicted individuals may be deterred from committing new crimes while incarcerated. On the other hand, if it has a limited effect, systemic barriers and social stigma in the labor market could force convicted individuals to remain in criminal activities. To gain further insight into these hypotheses, we investigate whether convicted individuals are more likely to re-offend. Our study reveals robust evidence that a criminal conviction leads to an increase in subsequent criminal activity, with a 13 percentage point higher probability of ever committing new crimes within three years compared to non-convicted counterparts. Similar results are observed for the intensive margin of recidivism.

Furthermore, we conduct a heterogeneity analysis by examining the instrumental variable estimates across various dimensions, including the severity of the crime, prior employment status, gender, and age. Our results indicate that the adverse effects on labor markets and the increase in subsequent criminal activity are found among individuals charged with low-severity crimes. These individuals are eligible for alternative penalties, such as fines, community services, and curfews, among other non-incarceration penalties. Among this group, the incapacitation effect plays a limited role in explaining the adverse employment outcomes, providing further evidence for the social stigma mechanism.

Our findings suggest that social stigma may be a potential mechanism driving the adverse effects of a conviction on labor market outcomes. Convicted individuals may face systematic obstacles in securing employment opportunities, leading them to remain involved in criminal activities. These results have significant policy implications, emphasizing the importance of developing strategies that address the root causes of criminal behavior and promote successful reintegration into society.

Our research is closely linked to several studies that utilize quasi-random judge assignment to examine the impact of incarceration on multiple outcomes. For instance, (KLING, 2006) found no evidence of negative consequences of incarceration duration on employment or earnings in California and Florida, while (GREEN; WINIK, 2010) did not find any effect of incarceration on recidivism in the District of Columbia. Similarly, (LOEFFLER, 2013) reported no detectable impact of incarceration on recidivism and employment outcomes in Chicago. On the other hand, some studies have found significant negative effects of incarceration. For instance, (AIZER; JR, 2015) found that juvenile incarceration reduces high school completion rates and increases adult recidivism in Chicago. In addition, (MUELLER-SMITH, 2015) indicated that incarceration increases recidivism and worsens labor market outcomes in Texas. (BHULLER *et al.*, 2018b) found no impact of a father's incarceration on children, while (BHULLER *et al.*, 2018a) found a positive spillover effect of incarceration on criminal networks and brother networks in Norway. (HARDING *et al.*, 2018) found a negative effect of incarceration on employment in Michigan. In the context of Sweden, (DOBBIE *et al.*, 2019) estimate the effect of parental incarceration on children's outcomes. In Norway, (BHULLER *et al.*, 2020) found that incarceration discourages recidivism, particularly among individuals who participate in employment programs. (HUTTUNEN; KAILA; NIX, 2020) estimate the impact of three types of punishments (fines, probation, and prison) on defendants' recidivism and labor market outcomes in Finland and find mixed results. (ARTEAGA, 2021) and (NORRIS; PECENCO; WEAVER, 2021) estimates the effect of parental incarceration on children and found beneficial effects on some children's outcomes in Colombia and US, respectively. Our study contributes significantly to this line of research as the majority of the studies have focused on the U.S. and Nordic countries. Using two quasi-experimental research designs, including random judge assignment, (GARIN *et al.*, 2022) find that incarceration has no long-term effect on labor market outcomes in US. In contrast, our paper provides the first set of causal evidence of the direct effects of a criminal conviction on labor and recidivism outcomes in a non-developed country.

Brazil's importance as the largest country in Latin America and the third-largest[2] prisoner population globally makes it a significant and relatively unexplored context for investigating the repercussions of criminal records in a high-crime environment where the State's control over its territory may be limited. The fact that our study considers such a context enhances its external validity, making the findings more applicable and relevant to similar settings in other regions with comparable challenges.

The structure of our paper is as follows. In Section 1.2, we provide contextual information on the Brazilian court system and explain the criminal case process. Section 1.3 outlines our research design. We describe our data and sample selection process in Section 1.4. Our main results for labor and recidivism, as well as our heterogeneity and robustness analyses, are presented in Section 1.5. Finally, we conclude in Section 1.6.

1.2 INSTITUTIONAL BACKGROUND

In this section, I discuss the main characteristics of the Judiciary System in Brazil and how it is composed. Also, I review how a criminal case starts and the institutional rules that must be followed.

1.2.1 Judiciary System in Brazil

The Judiciary System is one of the tripartite branches[3] that constitutes the Brazilian State. The national Constitution has organized the Judicial branch into the Common Justice and the Specialized Justice, each one with distinct competencies.

The Specialized Justice is composed of the Military Justice[4], Electoral Justice[5] and Labor Justice. Common Justice acts on all conflicts, but those within the sphere of Specialized Justice.

Within this aspect of residual justice, Common Justice is organized into two jurisdictional competencies: the Federal Justice, which operates within issues that involve Federal Union, and political crimes, among others; and the State Justice, which is responsible for all the remaining matters not due to any of the previously judicial branches.

[2] The US has the largest population, with more than 2 million prisoners, followed by China (1.69 million) and Brazil (811,000). World Prison Brief (2021), available at www.prisonstudies.org

[3] The Brazilian State is constituted by the Executive, Legislative, and Judiciary branches.

[4] Military Justice is responsible for prosecuting any military-related crime under the Brazilian Military Penal Code

[5] Electoral Justice acts on all electoral-related conflicts, as well as investigates electoral advertising, crimes, and any electoral process.

1.2.1.1 State Justice

By exclusion, the matters that are not under the competence of the Specialized and Federal Justice are under the responsibility of State Justice. The list of subject matters include civil, criminal, administrative law, just to cite a few.

The State Justice is organized according to the number of federal units (states) that compose the Federative Republic of Brazil and composed by two degrees of jurisdiction.

In the first degree, Trial Courtrooms are where any case starts, in which a decision is issued by a first instance judge. Any disagreement with the decision at the Trial Courtroom can be appealed to the States' Justice Courts (TJ), the second degree. The TJs are not trial courts. They only review the cases and then make a judgement to concur or dissent with the first instance decision. Second instance decisions can be appealed to the Superior Court of Justice (STJ). STJ is the highest appellate instance in the Brazilian Justice System. It is responsible for making the final decision on civil and criminal cases that do not involve constitutional matters. Seldom, decisions on cases that involve misunderstanding of the law or constitutional matters can be appealed to the Supreme Court (STF). These are exceptional cases.

The State Justice is territorially structured by Judicial Districts (*Comarcas in Portuguese*), Judicial Courts (*Foros in Portuguese*), and Trial Courtrooms (*Varas in Portuguese*). Judicial districts are units where first-instance judges extend their jurisdiction. They can be composed of one or more contiguous municipalities. Within Judicial districts, there can be one or more Judicial courts, which represent the physical space (buildings) where hearings are performed. They are called judicial courts because, in large enough judicial districts, they can have jurisdiction over specific portions of the territory. Finally, within Judicial courts lie one or more Trial courtrooms, the place that corresponds to one first-instance judge.

1.2.1.2 How does a criminal case work in Brazil?

The criminal justice process in Brazil are due to the Common Justice branch (mainly State Justice) and follows a set of laws and procedures established by the Brazilian Criminal Procedure Code (*Código de Processo Penal, in Portuguese*). In general, the process begins when someone files a criminal complaint or accusation against another person for an alleged crime. The main steps of the criminal process in Brazil are described below:

- **Preliminary investigation**: The preliminary investigation is conducted by the police and the Public Prosecutor's Office to collect evidence and determine if there is sufficient evidence to initiate a criminal action against the suspect. This phase may include obtaining statements, documents, and physical evidence.

- **Indictment**: If the preliminary investigation reveals evidence of a crime, the Public Prosecutor's Office files an indictment with the judge, who decides whether or not to accept the accusation. If the indictment is accepted, the criminal case is initiated.
- **Instruction phase**: During this phase, the judge may hear witnesses, request expert opinions, and examine the evidence presented by the parties. It is during this phase that the defendant is formally notified of the accusation and has the opportunity to defend themselves. The Public Prosecutor's Office can also present new evidence and witnesses.
- **Decision**: After the instruction phase, the judge may decide on one of the following options: acquit the defendant, convict the defendant, or partially acquit and partially convict the defendant. If the defendant is convicted, the sentence may include imprisonment, fines, community service, or other punitive measures.
- **Appeals**: Both the defendant and the Public Prosecutor's Office can appeal the judge's decision. Appeals are filed with the Provinces' Justice Courts (TJ) or the Superior Court of Justice (STJ), in the case of State Justice; and Federal Regional Court (TRF) or the Superior Federal Court (STF), in the case of Federal Justice.
- **Sentence enforcement**: If the conviction is upheld at all levels, the defendant must serve the sentence determined by the judgment. The sentence may include imprisonment in closed, semi-open, or open conditions, as well as other punitive measures determined by the judge.

The distribution of criminal cases in the Brazilian justice system is carried out through an electronic lottery system. This system is used both in State and Federal Justice.

The electronic lottery is a method of distribution that uses software to randomly select the courtroom that will be responsible for the case. This distribution method is important to ensure that cases are distributed fairly and without any external interference. The distribution process begins when the case is filed with the competent Judicial Court. Next, a unique case number (*Número Processual Único, in Portuguese*) is generated and registered in the electronic distribution system.

The criminal process is assigned to a courtroom, which may be composed of one or more judges, depending on the size and demand of the Judicial Districts. Generally, in courtrooms with a single judge, the case is automatically assigned to that judge after the assignment of the courtroom. In the case of courtrooms with more than one judge, the case is internally assigned among the magistrates, following criteria established by the judges themselves. Internal distribution may be carried out through predefined rules, such as the order of seniority of judges, the equitable distribution of cases among them, or through electronic lottery within the courtroom.

It's important to highlight that the distribution of cases is done randomly, aiming to ensure impartiality and neutrality in the judgment, without favoring or harming any of the parties involved in the process.

1.3 RESEARCH DESIGN

In order to estimate the effect of conviction on labor outcomes and subsequent recidivism, consider the model that relates future outcomes to an indicator of conviction:

$$Y_{i,t} = \beta_t \mathrm{I}_i + X_i^{\prime}\gamma + e_{i,t}, \tag{1.1}$$

where i denotes individual, t is the time of observation, β_t is the causal effect of interest, I_i is an indicator equal to 1 if defendant i is convicted, X_i is a vector of control variables, $Y_{i,t}$ is the outcome of interest measured t periods after case starts and $e_{i,t}$ is the error term. The problem of estimating Equation 1.1 is that any causal interpretation of β_t will be biased if conviction status is somehow correlated to any unobservable determinant of Y.

We address this endogeneity problem by exploiting the fact that criminal cases in Brazil are randomly assigned to courtrooms that differ systematically in their tendency of convicting (some courtrooms are more lenient than others). This feature leads to a random variation in the probability of being convicted that depends on the courtroom a defendant will be assigned.

Formally, we identify the causal impact of a conviction on defendants β_t using a measure of courtroom stringency (z) as an instrumental variable for being convicted. Our main specification is based on two-stage least squares (2SLS) estimation of β_t with the following two-equations system:

$$I_i = \delta z_{c,i} + X_i^{\prime}\theta + \varepsilon_i, \tag{1.2}$$

$$Y_{i,t} = \beta_t I_i + X_i^{\prime}\gamma + e_{i,t}, \tag{1.3}$$

where $z_{c,i}$ is our measure of stringency of the courtroom c assigned to defendant i's case and X_i is a vector of control variables for defendant i, including court-subject-year fixed effects representing the level at which randomization of courtrooms occurs.

Assuming the exogeneity and monotonicity of the instrument, the parameter β_t in Equation 1.3 can be interpreted as the local average treatment effect (LATE) of conviction for defendants who would have received a different decision if their case had been assigned to a different courtroom.

In line with the standard practice in research on judge fixed effects, we generate our instrument by utilizing the courtroom's inclination to convict in other cases, which helps to eliminate any correlation between the courtroom's ruling in a specific case and the value of the instrument. For each defendant i, we construct a measure of stringency of the initial courtroom his cases was assigned and use it as an instrument for being convicted. Following previous literature [(JR, 2008; TELLA; SCHARGRODSKY, 2013; MAESTAS; MULLEN; STRAND, 2013; DAHL; KOSTØL; MOGSTAD, 2014; FRENCH; SONG, 2014; AIZER; JR, 2015; DOBBIE; SONG, 2015; DOBBIE; GOLDIN; YANG, 2018; COHEN; YANG, 2019; BHULLER *et al.*, 2020; ARTEAGA, 2021; BHULLER; KHOURY; LøKEN, 2021; NORRIS; PECENCO; WEAVER, 2021; COLLINSON *et al.*, 2022)], we define the instrument as the difference of two leave-one-out average of the conviction indicator:

$$Z_{f,c,s,i} = \frac{1}{(n_{f,c,s}-1)}\left(\sum_{k=1}^{n_{f,c,s}}(I_k) - I_i\right) - \frac{1}{(n_{f,s}-1)}\left(\sum_{k=1}^{n_{f,s}}(I_k) - I_i\right), \tag{1.4}$$

where i indexes defendants, f courts, c courtrooms and s refers to subject matter. The variable I is an indicator equal to 1 if defendant i is convicted, $n_{f,c,s}$ is the number of cases of subject s in court f and courtroom c and $n_{f,s}$ is the number of cases of subject s in court f. This instrument can be interpreted as a measure of how stringent one courtroom is compared to the court it belongs to when ruling a certain type of criminal case. One advantage of this instrument is that it captures the level of leniency of a courtroom within the same pool. When using this measure, we always condition on fully interacted court-subject-year fixed effects to account for the fact that randomization occurs within the same pool of courtrooms. This guarantees we are limiting the comparison of defendants on the verge to be assigned to the same set of courtrooms.

Figure 1.1 presents the first results of our instrument and it illustrates a considerable variation. The histogram shows that a courtroom at the 90^{th} percentile in the distribution convicted around 62% of defendants, compared to 32% for a courtroom at the 10^{th} percentile. The average courtroom stringency rate is 47% with a standard deviation of 5%.

Moreover, in the presence of heterogeneous effects, one concern is whether the assumption of monotonicity holds, meaning that a defendant who would be convicted by a less stringent courtroom would also be convicted by a stricter courtroom, and vice versa for non-conviction. To address this issue, we conducted two sets of tests in Section 1.5, both of which suggest that monotonicity is likely to hold. Additionally, another concern is about how we create our measurement of courtroom stringency. In our main specifications, we measured courtroom stringency as the leave-one-out mean conviction rate, which averages the conviction rates in other cases a courtroom has handled while excluding the case being studied. Following (BHULLER *et al.*, 2020) and (NORRIS; PECENCO; WEAVER, 2021), we perform tests using alternative measures of $Z_{f,c,s,i}$ and a split-sample approach. Overall, our results provide support for the validity of our research design.

Furthermore, in order to address any potential serial correlation among defendants at the randomization level, we follow (CHAISEMARTIN; RAMIREZ-CUELLAR, 2023) and employ a clustering approach for the standard errors in both the first and second stages, with clustering at the court levels.

1.4 DATA

In order to estimate the impact of conviction on labor and recidivism outcomes, we performed a unique merge between individual criminal cases and a rich set of administrative data in Brazil.

Data on criminal cases in Brazil was gathered from two sources. The primary source was text sentences from all criminal adjudicated cases filed at the State Court of São Paulo (*Tribunal de Justiça de São Paulo, TJSP*)[6]. This dataset covers the period between 2010 and 2022, consisting of more than 1.7 million sentenced cases. The dataset includes a unique identifier for each case, as well as information on the district, court, courtroom, judge, subject matter, and text of the sentence. The second source of information was proprietary data from a private firm that collects judicial data from multiple Brazilian courts. This information includes the text of the sentence (when the case is sentenced), the names of plaintiffs and defendants, and whether the case was randomly assigned to a courtroom. This dataset covers the period between 2010

[6] TJSP is the largest court in Brazil and handles over a quarter of the country's judicial proceedings.

and 2022 and handles over 30 million adjudicated and pending cases. By combining these two datasets, a comprehensive picture of criminal cases in Brazil was obtained, which allows us both to measure the treatment variable (convicted or not) as well as to track future criminal behavior.

To obtain the final decision from the text data of the sentences, algorithms based on regular expressions were developed. These algorithms are designed to code the conviction decision from the text of the sentence. The process of extracting the decisions involves two steps. First, the algorithms identify which text from each case pertains to the convicted/not convicted decision. Second, once the relevant text has been identified, the algorithms extract the sentence containing the decision. This allows for accurate and efficient mining of the decisions from the text data of the sentences. Overall, we are able to retrieve 2,535,674 criminal case decisions from 2010-2022 period.

In order to perform a fined merge with other administrative data sources, we augmented our criminal case dataset with individual identification information from the *Cadastro de Pessoas Físicas (CPF)* registry, a database also provided by the previous firm. This comprehensive database covers almost the entire Brazilian population and provides unique identifiers for each individual, along with other important information such as their birth date, gender, and mother's name. Enhancing our criminal case with such unique identifiers will allow us to perform further merges with other administrative data sources.

Finally, our study employs the *Relação Anual de Informações Sociais (RAIS)*, a dataset that encompasses all formal workers and firms in Brazil from 2002 to 2020. This extensive dataset provides crucial information, including job start and end dates, job location, unique identities of employers and employees, contract type, occupation and sectoral codes, worker education, race, earnings, and many others. With access to this dataset, we construct measures of labor outcomes such as the yearly total number of days worked and total earnings. Utilizing this data allows us to comprehensively analyze and evaluate labor market outcomes in Brazil.

1.4.1 Matching

Our study is faced with the significant challenge of linking defendants from criminal cases to various data sources. To tackle this issue, we implemented a rigorous and systematic approach.

Firstly, we utilized algorithms based on regular expressions to extract defendant names from the text sentences collected from *TJSP* criminal cases, resulting in 834,261 defendant names being retrieved. To expand this dataset, we partnered with a private firm that specializes in collecting judicial data from multiple Brazilian courts, allowing us to obtain additional criminal cases from other State Courts, which already included the defendants' names. This procedure resulted in approximately 10 million names being retrieved, significantly enhancing the scope and depth of our dataset.

Secondly, we enhanced our compiled criminal case dataset by incorporating individual identification information from the *Cadastro de Pessoas Físicas (CPF)* registry, which provides unique identifiers for each individual, along with other important information such as birth date, gender, and mother's name. We leveraged this dataset by assigning *CPF* to defendant names that we found to be unique in this registry. With this procedure, we were able to assign unique identifiers to around 50% of the defendant names, enabling us to accurately link defendants to multiple data sources, such as RAIS, which also presents such identifiers. Overall, our approach allowed us to create a robust and comprehensive dataset for our study.

1.4.2 Sample Selection

The sample for this study comprises criminal cases where a sentence was issued between 2010 and 2022, totaling 2,814,081 cases. However, certain restrictions were applied to refine the dataset. Firstly, cases that were not randomly assigned to a courtroom were excluded. Removing non-randomly assigned cases from the dataset is a simple process, as we are able to identify and flag them in our records. These cases were either assigned to specific courtrooms due to their connection with other cases or because of judicial rules that mandate certain courtrooms to rule on specific cases. Secondly, the dataset was limited to courtrooms that had at least 10 cases per year and subject matter during the period. We make this restriction in order to avoid noise when calculating our instrument. Additionally, to enhance the precision of our estimates, we incorporated court-by-subject matter-by-year fixed effects into our analysis. Consequently, we only considered cases from courts that had a minimum of two courtrooms receiving cases from a particular subject matter in a given year. As a result of these restrictions, a sample of 579,684 randomly assigned cases was obtained, all of which were assigned to the same pool of courts that had at least two courtrooms ruling on at least 10 cases per subject-year. This refined sample was used to construct the instrument variable for the study.

For our estimation sample, we further restrict our sample to defendants with age between 25-55 at the start of their cases and whose labor outcomes can be linked anytime between 2002-2020. In addition, we limit our analysis to criminal cases that started between 2010 and 2017 period. This duration ensures that each defendant can be tracked and followed for, at least, five years before up to three years after the case filing, providing a more comprehensive understanding of the potential effects of an conviction on labor and criminal behavior outcomes. After applying these restrictions, our baseline estimation sample comprises 41,646 cases, involving 42,597 defendants, across 961 courtrooms.

1.4.3 Descriptive Statistics

Table 1.1 offers a comprehensive overview of the defendant characteristics in our baseline sample, shedding light on the demographic, socioeconomic, and employment characteristics of individuals involved in the criminal justice system in Brazil during the 2010-2020 period.

[Tabela 1.1 about here.]

Column (1) of Table 1.1 presents the descriptive statistics of all the individuals included in our analysis. The results reveal that the vast majority of defendants are male, representing around 86% of all individuals, while females account for 14% of the sample. The average age of defendants at the time the cases are filed is 35 years old, with a predominance of White individuals with at least a high school education (12 years of education or more). Furthermore, roughly half of the sample had a job in the year before the case was filed, with over 65% employed in the years prior.

Columns (2) and (3) of Table 1.1 enable us to delve deeper into the characteristics of defendants who were convicted versus those who were not convicted. The findings suggest that convicted defendants are negatively selected across several variables, including race, education, and prior employment status. Specifically, convicted defendants tend to be composed of a higher share of Black individuals (29% versus 24% for the not convicted group) and a lower share of Whites (68% versus 73% for the not convicted group). Additionally, convicted defendants tend to be less educated and younger than their not convicted counterparts, and have significantly worse employment status in the years leading up to the criminal charge, with only 46 percent of them working in the previous year against almost 60% from not convicted ones.

1.5 MAIN RESULTS

1.5.1 Instrument Validity

Some conditions are necessary to interpret our estimation of β_t in Equation 1.3 as the local average treatment effect of conviction on labor and recidivism outcomes.

The first of these conditions is the instrument relevance condition, which requires that the instrument used in the analysis must be correlated with the conviction decision. Figure 1.1 provides a visual representation of this condition. The histogram illustrates the wide variation in our instrument, with courtroom stringency rates ranging from 0.32 to 0.62 across courtrooms at the 1% and 99% percentile at the distribution, respectively, with a mean of 0.47 and a standard deviation of 0.05. The fitted line on the graph depicts the estimates obtained from a local regression of the conviction decision as a function of courtroom stringency, revealing a strong first stage. Specifically, as courtroom stringency increases, conviction rates also increase, suggesting a significant correlation between our instrument and the conviction decision.

[Figura 1.1 about here.]

Table 1.2 provides further insight into the strength of our instrument by presenting the results of our first stage equation. The findings indicate a robust and highly significant relationship between our instrument and the conviction status, showing that assignment to a courtroom with a 10 percentage point higher probability of conviction leads to an 8 percentage point increase in the likelihood of being convicted. Given the conviction average rate of 0.62, this result represents a 13% deviation from the mean. Our findings are robust to the inclusion of various controls and adjustments to the fixed effects formulation.

[Tabela 1.2 about here.]

Second, we also need our instrumental variable not to be correlated with both defendant and case characteristics that could influence the defendant's subsequent outcomes. This is called the exogeneity assumption. Table 1.2 provided the first set of evidence for this assumption. If criminal cases are randomly assigned to courtrooms, then adding controls should not influence the estimates of the first stage. As we can see, extending the number of controls and changing the fixed effects formulation does not substantially affect the coefficient. To further support the exogeneity assumption, we conducted a direct test for random assignment by investigating whether the defendant and case characteristics could explain the allocation of criminal cases to courtrooms. In Table 1.3, the first column presents a regression of conviction outcome on the relevant covariates, while the second column shows a regression of courtroom stringency on the same set of characteristics. The results reveal that while the case and defendant characteristics are highly predictive of the criminal conviction indicator, they do not have any noticeable effect on courtroom stringency, providing empirical support for the random assignment of our instrumental variable. Although we find a statistically significant result for the *Female* indicator in the second column of Table 1.3, the coefficient is particularly small and does not represent a meaningful result. Thus, we do not reject the null hypothesis for the joint test for significance at 10%.

[Tabela 1.3 about here.]

Third, as we assume the effect of conviction on the subsequent outcomes to differ across individuals, we need the monotonicity assumption to hold. In our setting, monotonicity means that if a lenient courtroom convicts a defendant, a more stringent would also convict (and vice versa for non-conviction). This is called the no-defier assumption. With this assumption, it is possible to interpret the β_t as a local average treatment effect. In other words, the estimated effect represents the average causal effect among a specific subgroup of defendants who would have potentially received a different conviction decision if their case had been assigned to a different courtroom. One implication of this assumption is that the first-stage estimates should be non-negative for any subsample. Following (BHULLER *et al.*, 2020) and (NORRIS; PECENCO; WEAVER, 2021), we conducted two sets of tests. First, we perform first-stage estimations on different subsamples of the data, including quartiles of a constructed index of all the characteristics used in Table 1.1, previous employment status, education, age, and race. Second, we employed a reverse-sample testing method, dividing the data into the same subsets as the first test, but redefining the instrument for each subset as the conviction rate of cases that were not part of that estimation subset. Our results, reported in Table 1.4, confirm that the coefficient on courtroom stringency remained consistent in sign across all subsets, thereby providing evidence for the validity of the monotonicity assumption.

[Tabela 1.4 about here.]

1.5.2 Effects of conviction on labor

This study aimed to investigate the impact of criminal convictions on labor outcomes in Brazil, with a focus on employment and earnings. By analyzing the extensive and intensive margins of employment, as well as total earnings, we aimed to provide a comprehensive understanding of the impact of criminal convictions on individuals' labor market outcomes. Figure 1.2 shows the IV estimates of the effects of a conviction on labor outcomes in a given year. Each point on the graph is the β_t coefficient from period-by-period versions of our 2SLS equations. In Table 1.5, we summarize these results while adding more elements to our analysis.

[Figura 1.2 about here.]

Figure 1.2a shows the IV estimates of the effects of a conviction on the extensive margin of employment in a given year. We define being formally employed as if the defendant worked for some period in a given year. Our results indicate that, on average, a conviction leads to a substantial decrease in the probability of being formally employed in Brazil in the first year after the case starts, and this negative impact persists for up to three years following the case filing. Importantly, we also find that the difference in the probability of being employed between convicted and non-convicted individuals prior to the filing is not statistically significant. This finding provides more evidence for the validity of our research design, as it indicates that our instrument is not correlated with previous labor outcomes. We also investigated the cumulative effect of criminal convictions on employment outcomes, as shown in Figure 1.2b. Our results demonstrate a declining trend in the probability of obtaining formal employment over time for convicted individuals. This indicates that the negative impact of criminal convictions on labor outcomes may extend beyond the immediate aftermath of a case filing. In Table 1.5, columns (1)-(2) summarize these results, showing that the probability of being employed in any given year (up to 3 years after case filing) reduces by 8 percentage points, equivalent to a sizable 18% drop from the average, while the probability of ever working within 3 years reduces by 12 percentage points, representing a nearly 22% fall from the average.

[Tabela 1.5 about here.]

In addition to examining the impact of criminal convictions on the extensive margin of employment, we also investigated the effect on the intensive margin. Specifically, we measured the intensive margin as the total number of days formally employed in a given year. To facilitate interpretation, we follow (NORRIS; PECENCO; WEAVER, 2021) and transform the intensive margin outcomes using the inverse hyperbolic sine (I.H.S) function, allowing for the interpretation of the results as percent changes. Our analysis, depicted in Figure 1.2c, reveals that criminal convictions lead to a significant and persistent reduction in the number of days formally employed, lasting for up to three years after the case filing. Moreover, the cumulative effect of criminal convictions on the total number of days worked, as illustrated in Figure 1.2d, indicates a worsening trend in the labor outcomes of convicted individuals. This underscores the enduring impact of criminal convictions on employment outcomes, demonstrating the long-lasting negative consequences that these convictions can have on individuals' labor market prospects. These results are summarized in columns (3)-(4) of Table 1.5. It shows that the average number of days worked in any given year (up to 3 years after case filing) reduces by 64%, while the cumulative measure reduces by 77%.

Finally, we analyzed the effect of criminal convictions on earnings, as shown in Figures 1.2e and 1.2f. We measured earnings (expressed in units of thousands) as the sum of all (real) salaries[7] received by the defendant in the year. We found that total earnings were significantly reduced following a case filing, indicating the long-lasting and severe impact of criminal convictions on individuals' earnings prospects. Moreover, the cumulative effect of criminal convictions on earnings, as shown in Panel 1.2f, highlights the persistently negative impact of criminal convictions on individuals' earnings, which worsens over time. In Table 1.5, columns (5)-(6) show that earnings (up to 3 years after case filing) reduce by 13%, while the cumulative measure reduces by 26%.

Overall, the results of our study provide strong evidence of the detrimental effects of criminal convictions on labor outcomes in Brazil. These findings have significant policy implications, emphasizing the urgent need for effective interventions to alleviate the negative consequences of criminal records on individuals' labor market prospects.

1.5.3 Effects of conviction on recidivism

The previous analysis indicates that individuals who are marginally convicted are likely to experience adverse effects on their labor market outcomes. However, the underlying mechanisms responsible for this phenomenon remain uncertain. One possible explanation is that the social stigma associated with a criminal record hinders these individuals from securing

[7] We calculate our measure of real salary by adjusting the nominal values for inflation using the Extended National Consumer Price Index (*IPCA, in Portuguese*). The salaries are measured at constant prices as of 2020.

employment opportunities. Alternatively, the incapacitation of convicted individuals, who are more likely to be incarcerated, could limit their employment prospects. Determining the contribution of incapacitation is challenging since we do not have access to the prison time of a particular conviction. If incapacitation plays a significant role, convicted individuals may be deterred from committing new crimes while incarcerated. However, if incapacitation has a limited effect, systemic barriers and stigma in the labor market may force convicted individuals to continue in criminal activities.

To shed further light on these hypotheses, we examine whether convicted individuals are more likely to reoffend. Our analysis includes two measures of recidivism outcomes: the probability of being charged with at least one new crime by the end of a post-filing year (reflecting the extensive margin of reoffending), and the cumulative number of new criminal charges by the end of a post-filing year (reflecting the intensive margin of reoffending). We conduct separate estimations for each measure based on the severity. We categorize the severity of a criminal case as either *severe* or *non-severe* based on their base-penalty exceeding (or not) 4 years of sentence[8]. Figure 1.3 and Table 1.6 present the IV estimates of the effect of conviction on criminal recidivism.

[Figura 1.3 about here.]

As shown in Figure 1.3a, the probability of reoffending increases over time. Within the first few years after filing, the likelihood of a convicted individual being charged with a new crime rises by nearly 5 percentage points. This negative effect persists throughout the 3-year period, reaching more than 10 percentage points. We also find increasing trends in the probability of committing new non-severe and severe crimes, based on the breakdown of results by case severity.

Similarly, the results for the intensive margin of recidivism, presented in Figure 1.3b, also show a steady and increasing effect of conviction on the total number of criminal charges. For instance, within three years after filing, the total number of reoffenses during that period is almost 13% higher for convicted individuals. We also observe a similar pattern for the total number of new charges by case severity, although the results for severe cases are imprecise.

[8] This classification is based on Article 44 of the Brazilian Penal Code (*Código Penal, in Portuguese*). Specifically, cases with a base-penalty exceeding 4 years of sentence are classified as *severe* because they do not qualify for alternative sentencing options such as fines, community service, curfews, or other non-incarceration penalties. On the other hand, *non-severe* subjects are the ones with a base-penalty of less than 4 years of sentence, and are usually exchanged with non-incarceration options.

Table 1.6 summarizes all of the recidivism results. For all measures, except for severe charges, we found a significant and substantial effect of conviction on recidivism. The probability of ever being charged with any, non-severe, and severe crime within 3 years after case filing increases by 12.7, 5.1, and 2.8 percentage points, respectively. We also found a similar effect of conviction on the intensive margin of recidivism, with the cumulative number of any, non-severe, and severe charges within 3 years increasing by 13.1, 5.1, and 2 percentage points, respectively, although the results for severe charges are not statistically significant.

Overall, our findings suggest that social stigma might play a significant role in the adverse effects of conviction on labor market outcomes. Convicted individuals may face systemic barriers in securing employment opportunities, which could lead them to remain involved in criminal activities.

[Tabela 1.6 about here.]

1.5.4 Heterogeneity

In the main analysis, we estimate the local average treatment effect of conviction on labor and recidivism outcomes. Our results indicate that individuals who are marginally convicted are likely to experience adverse effects on their labor market outcomes, which could lead them to remain involved in criminal activities. These findings emphasize the importance of developing policies aimed at tackling the root causes of criminal behavior, promoting successful reintegration into society, and reducing recidivism.

To further explore the effects of conviction on these outcomes, we conduct a heterogeneity analysis by examining the IV estimates across multiple dimensions, including the severity of the crime, previous employment status, gender, and age. Our results, presented in Table 1.7, provide insights into how the impact of conviction may vary based on these factors.

[Tabela 1.7 about here.]

Crime Level. The results presented in Table 1.7 shed light on the effects on labor and recidivism by crime severity, both on the extensive and intensive margins. Our findings, shown in Panels A-F, indicate that individuals convicted for severe crimes face significantly lower employment and earnings levels, although results for those convicted of non-severe crimes are also negative and statistically significant. Specifically, the probability of ever working within 3 years after case filing drops by 19 percentage points for individuals convicted for high-severity crimes, while it drops by 10.5 percentage points for those convicted for low-severity crimes. Similarly, the results for earnings show declines of 28% and 10% for the severe and non-severe groups, respectively.

In terms of recidivism, we observe that the effect is concentrated in individuals convicted for non-severe crimes. The main effect of the extensive and intensive margin of recidivism (Panel F and H, respectively) is 1 and -3 percentage points, respectively, compared to 15 and 17 percentage points for those convicted for low-severity crimes. Moreover, examining recidivism of severe and non-severe cases reveal that individuals convicted for non-severe crimes also display larger effects in both margins of recidivism. Specifically, Panels I-M show the probability of ever committing and the total number of new non-severe and severe crimes are much larger for the non-severe convicted group.

It is noteworthy that analyzing labor and recidivism results by crime level is not only critical for exploring heterogeneous effects but also for providing further evidence of the potential social stigma mechanism. Given that convictions for non-severe crimes are eligible for non-incarceration penalties, the incapacitation effect has a very limited role in explaining the adverse employment results.

Previous Employment Status. Table 1.7 also sheds light on the impact of defendants' previous employment status on the relationship between conviction, employment, and recidivism outcomes. Columns (4) and (5) present estimates by the defendants' previous employment status.

The results suggest that previous employment status does not substantially alter the adverse impact of conviction on labor market outcomes. Panels A-F shows that both previously employed and unemployed individuals experience significant decreases in employment and earnings levels following a conviction.

Moreover, our analysis reveals little heterogeneity in the effect of conviction on recidivism across the two groups, except for severe crimes. In Panel L-M, We show a substantial difference between the previously employed and unemployed groups for severe crimes in both the extensive and intensive margins. Specifically, the probability of ever committing and the total number of new severe crimes within 3 years of the case filing increase by 5.7 percentage points and 4.6 percentage, respectively, for the previously employed convicted group, while the estimates for the previously unemployed group are almost negative.

In summary, despite potential differences in their pre-conviction employment status, both groups face adverse labor market outcomes following conviction. However, the impact of conviction on recidivism only substantially differs between the previously employed and unemployed groups for severe crimes.

Gender. We also investigate the potential heterogeneity in the effect of conviction by gender. Our results are presented in Columns (6) and (7) of Table 1.7. As expected, our sample is primarily composed of male defendants, which limits our ability to analyze gender differences.

Our main findings are concentrated on the male group, and we do not find any statistically significant results for the female group across all measures of labor. Similar results were found for recidivism outcomes. However, this may be due to a smaller sample size for this subgroup.

Interestingly, we observe that the results for recidivism (Panels G-M) in both margins have the opposite sign for the female group. This could potentially suggest a decrease in reoffending. In fact, Panel J shows that the number of cumulative new non-severe charges decreases by 30% within three years for females, and this result is statistically significant at 10% level. The other results show a similar pattern but we cannot rule out the possibility that this effect is zero due to the lack of statistical power.

Examining gender heterogeneity provides useful insights into the differential impact of conviction on men and women. However, given the limited sample size of females in our dataset, further research is needed to confirm the potential pattern that we found.

Age. Another interesting source of heterogeneity is age, as the effects of conviction might differ between younger and older individuals. To explore this, we divided the sample into those under and over the age of 35, and the results are presented in Columns (8) and (9) of Table 1.7.

Our findings suggest that the adverse effects of conviction on labor outcomes are concentrated in the younger group. Specifically, Panels A-H present that those under 35 years old experience a larger negative impact on their employment and earnings prospects compared to their older counterparts.

In terms of recidivism, we find that older convicted individuals have a larger effect on recidivism in general. Interestingly, when we delve into the different types of recidivism, we notice that older individuals are more inclined to commit low-severity crimes again, while younger individuals under the age of 35 are more likely to reoffend in more serious crimes.

The analysis of heterogeneous effects presented in Table 1.7 provides a nuanced understanding of the impacts of criminal conviction on labor and recidivism outcomes for different subgroups of defendants. The results suggest that the adverse effects of conviction are not uniform across all subgroups and may vary based on factors such as the severity of the crime, previous employment status, gender, and age.

Our findings indicate that individuals who have been convicted of non-severe crimes are more susceptible to re-offending than individuals convicted by more serious crimes. Additionally, both groups experience adverse effects on their labor outcomes. Furthermore, our research reveals that younger individuals experience more severe negative impacts on their employment and earnings prospects and are more likely to commit serious crimes than those who are over 35 years old. These findings highlight the importance of considering these distinct subgroups when designing interventions aimed at reducing the negative effects on individuals with a criminal record and improving their chances of successful reintegration into society.

1.5.5 Robustness

In order to ensure the robustness of our main findings, we conducted additional analyses using different criteria to calculate our instrument. Our results indicate that our conclusions are not dependent on the specific method used to construct it.

Table 1.8 presents the results of our analyses. The first column shows our baseline findings for comparison. Columns (2)-(5) depict the results when we used the leave-one-out conviction rate for courtrooms that handled at least 5, 15, 20, and 25 cases of a subject within a year, respectively. The estimated effects were consistent across all specifications. Panel A shows the results of our first stage, while Panels B-N present the results of our labor and recidivism outcomes.

[Tabela 1.8 about here.]

Furthermore, we randomly split our sample and used one part to calculate the average conviction rate of each courtroom, then used these measures of courtroom leniency as an instrument for conviction in the other part of the sample. The resulting estimates were similar to our baseline findings.

These results provide additional support for the reliability of our research design, as they demonstrate that our conclusions are not sensitive to the number of cases per courtroom. Overall, our findings remain robust across different specifications of the instrument used in our analysis.

1.6 CONCLUSION

The criminal justice system serves a critical role in society, and its impact on individuals who have been convicted of a crime cannot be overstated. While the immediate consequences of legal sanctions, such as imprisonment or fines, are well-documented, the long-term effects of a criminal record on an individual's life are less understood. Therefore, this paper represents an important step forward in our understanding of the impacts of criminal convictions by providing the first causal estimates of the effect of conviction on labor and recidivism outcomes in Brazil.

Our results indicate that individuals who are marginally convicted experience adverse effects on the labor market. Specifically, we found that they faced challenges in securing employment opportunities, which lead them to remain involved in criminal activities. There are various possible mechanisms that may mediate this effect, including the incapacitation effect and social stigma faced by convicted individuals in the labor market. However, we argue that the latter plays a more significant role, while the former has only a limited effect.

Despite the importance of our evidence, several questions remain open about the spillover effects of criminal convictions. One potential line of research is to understand the effect of such convictions on family dynamics. The results are particularly relevant from a policy perspective as the societal costs and benefits of conviction could be magnified or muted once these spillover effects are taken into account. Thus, additional research along these lines is needed to provide a more comprehensive understanding of the impact of criminal convictions on society.

Overall, the evidence presented in this paper underscores the need for policymakers to address the root causes of criminal behavior. Effective interventions that promote successful reintegration into society should be prioritized to prevent the cycle of crime and recidivism. Such policies will not only benefit the individuals who have been convicted but also society as a whole by reducing crime rates and improving public safety. Therefore, the findings of this paper have important implications for policymakers seeking to understand and address the complex issues related to criminal justice.

FIGURES

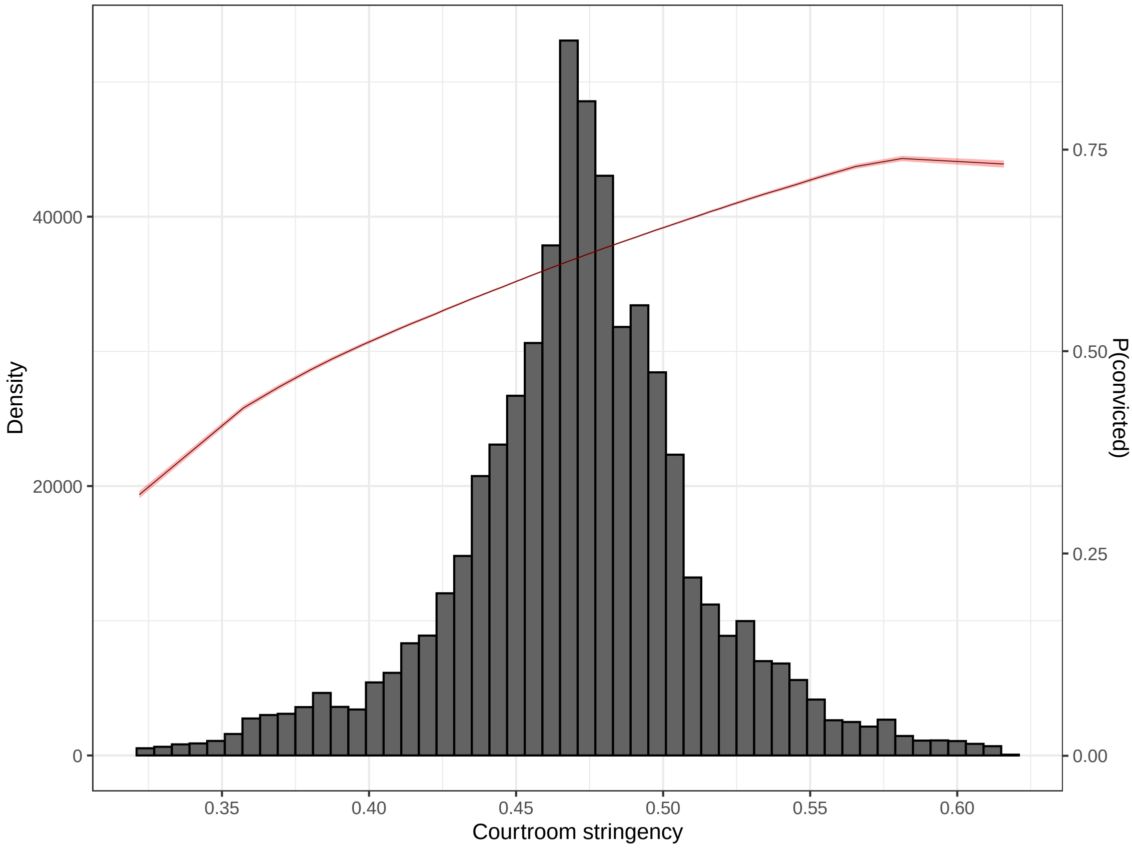
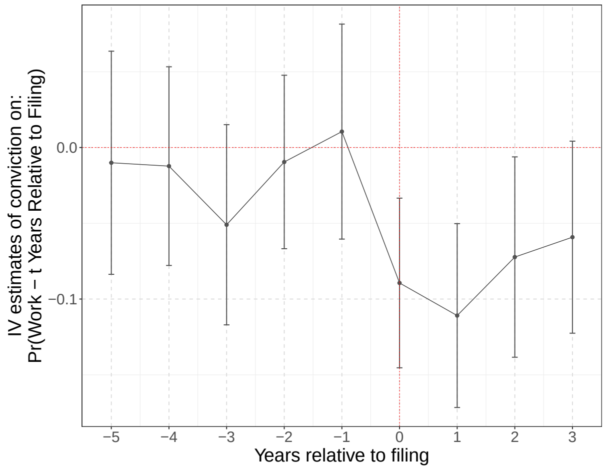
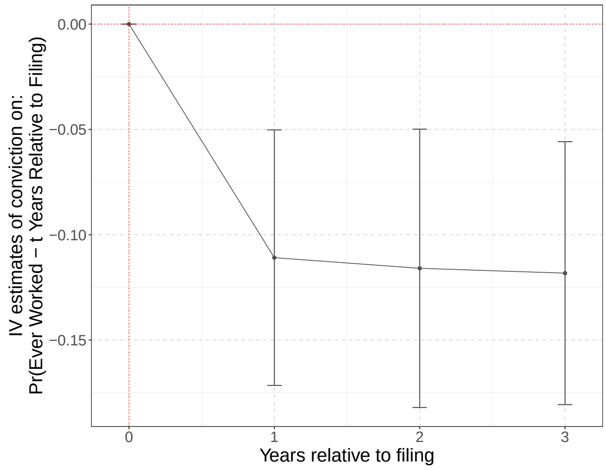


Figura 1.1 – First Stage Graph Of Conviction On Courtroom Stringency

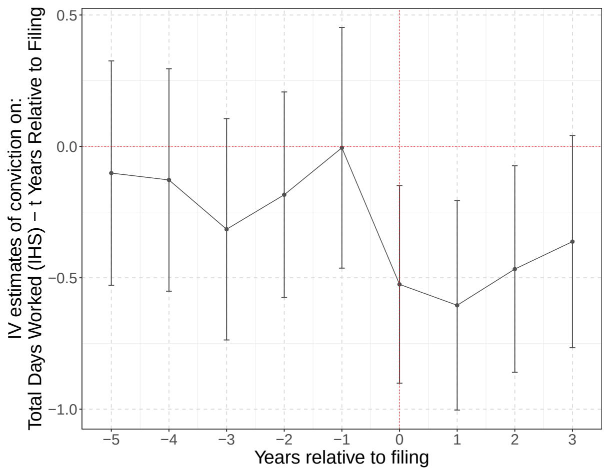
Notes: Baseline sample filed 2010-2020 in Brazil. The probability of conviction was plotted on the right y-axis against the leave-one-out average of courtroom stringency. The plotted values are mean-standardized residuals obtained from regressions on *court* × *year* × *subject* fixed effects. The solid line depicts a local linear regression of conviction on the instrument, while the dashed lines represented 95% confidence intervals. The histogram shows the density of courtroom stringency along the left y-axis, with the exclusion of the top and bottom 1%.
Source: Prepared by the author.



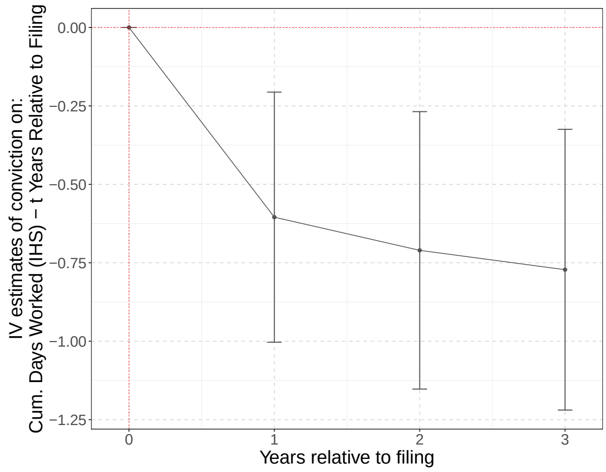
(a) Pr(Work)



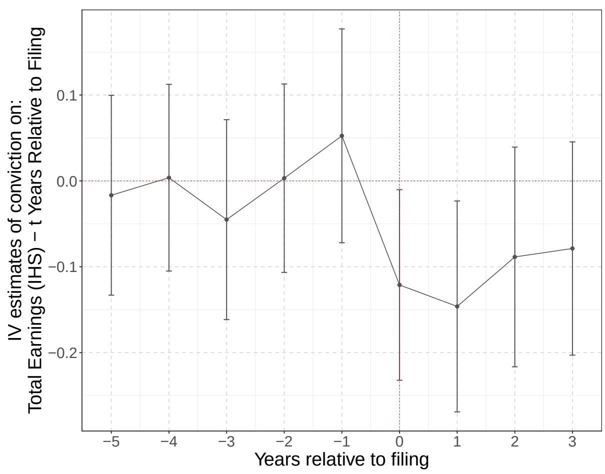
(b) Pr(Ever Worked)



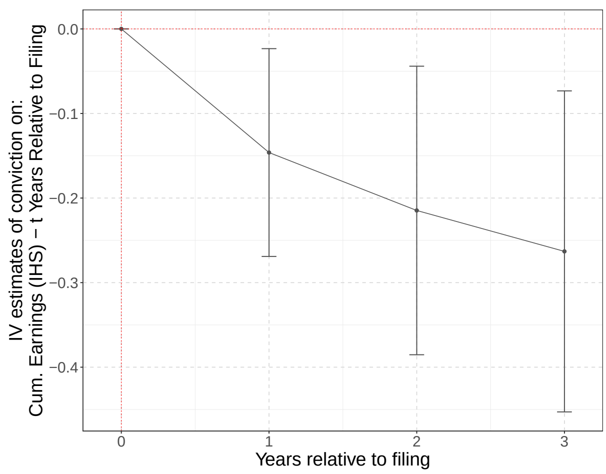
(c) I.H.S. Total Days Worked



(d) I.H.S. Cum. Total Days Worked



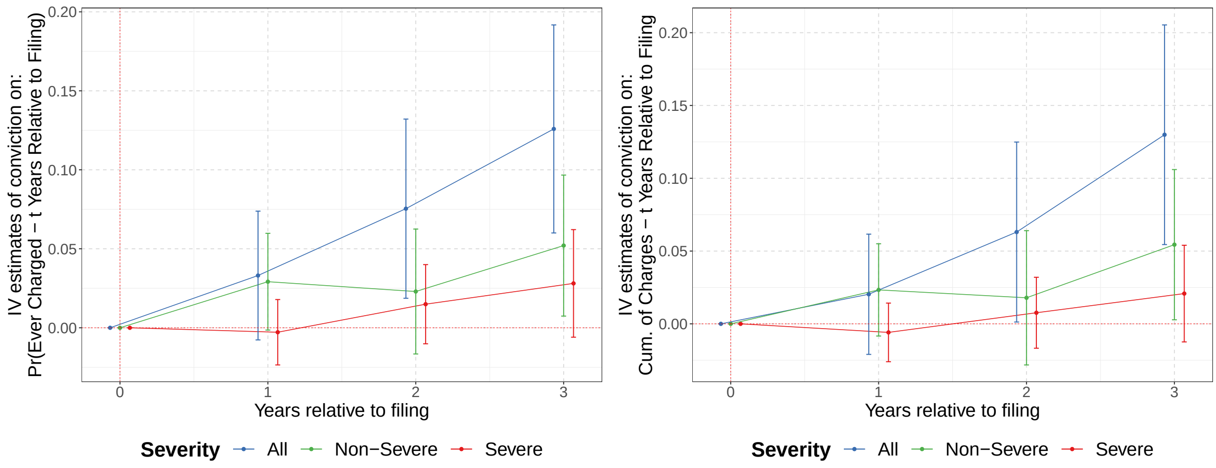
(e) I.H.S. Total Earnings (real)



(f) I.H.S. Cum. Total Earnings (real)

Figura 1.2 – The Effect of Conviction on Labor Outcomes

Notes: Baseline sample of criminal cases filed 2010-2017 in Brazil. I.H.S stands for *Inverse Hyperbolic Sine*. Total Earnings and Cum. Total Earnings are expressed in units of thousands of *Reais*. Each point on the graph is the IV estimation from period-by-period version of our 2SLS formulation. The error bars show 90% confidence intervals.
Source: Prepared by the author.



(a) Pr(Ever Charged) **(b) I.H.S. Cum. Number of New Charges**

Figura 1.3 – The Effect of Conviction on Recidivism

Notes: Baseline sample of criminal cases filed 2010-2017 in Brazil. I.H.S stands for *Inverse Hyperbolic Sine*. Each point on the graph is the IV estimation from period-by-period version of our 2SLS formulation. The error bars show 90% confidence intervals.

Source: Prepared by the author.

TABLES

Tabela 1.1 – Descriptive Statistics

	Overall (1)	Not convicted (2)	Convicted (3)
Gender			
Male	85.96%	83.17%	88.93%
Female	14.04%	16.83%	11.07%
Age	35.02 (7.81)	36.37 (8.15)	33.57 (7.15)
Race			
White	70.42%	72.87%	67.88%
Black	26.55%	24.26%	28.92%
Indigenous	0.10%	0.10%	0.10%
Non Identified	2.93%	2.76%	3.10%
School			
< 9 years	6.28%	6.30%	6.26%
< 12 years	16.53%	15.65%	17.75%
>= 12 years	77.19%	78.05%	76.00%
Employed, year t-1	0.53 (0.50)	0.59 (0.49)	0.46 (0.50)
Employed, year t-2 to t-3	0.66 (0.47)	0.70 (0.46)	0.62 (0.49)
Employed, year t-4 to t-5	0.69 (0.46)	0.72 (0.45)	0.66 (0.47)
Missing Xs	0.53 (0.53)	0.48 (0.53)	0.58 (0.52)
Observations	56,723	29,347	27,376

Notes: Baseline sample of criminal cases filed during 2010-2020 period. Statistics are at the defendant level and include $56,723$ unique defendants. Column (1) reports the sample averages/proportions for the full sample. Columns (2) and (3) reports the sample averages/proportions for the 'Not convicted' and 'Convicted' sub-sample, respectively. Standard deviations are displayed in parenthesis.
Source: Prepared by the author.

Tabela 1.2 – First stage estimates of conviction on courtroom stringency

	P(convicted)			
	(1)	(2)	(3)	(4)
Courtroom stringency	0.827***	0.822***	0.821***	0.821***
	(0.028)	(0.027)	(0.027)	(0.025)
Age		-0.003***	-0.003***	-0.003***
		(0.0004)	(0.0003)	(0.0006)
Female		-0.066***	-0.068***	-0.068***
		(0.006)	(0.006)	(0.006)
Black race		0.021***	0.017***	0.017*
		(0.007)	(0.006)	(0.009)
Indigenous race		0.054	0.042	0.042
		(0.049)	(0.049)	(0.039)
Non identified race		0.018	0.010	0.010
		(0.012)	(0.011)	(0.007)
<= 12 years education		-0.0010	-0.002	-0.002
		(0.013)	(0.013)	(0.006)
> 12 years education		-0.015	-0.022**	-0.022***
		(0.010)	(0.011)	(0.006)
Missing Xs		0.037***	-0.025***	-0.025***
		(0.006)	(0.008)	(0.008)
Worked, t-1			-0.071***	-0.071***
			(0.009)	(0.006)
Worked, t-2 to t-3			-0.001	-0.001
			(0.004)	(0.001)
Worked, t-4 to t-5			-0.014***	-0.014***
			(0.004)	(0.003)
Court-Year-Subject FE	Yes	Yes	Yes	No
Court FE	No	No	No	Yes
Year FE	No	No	No	Yes
Subject FE	No	No	No	Yes
Dependant mean	0.617	0.617	0.617	0.617
Observations	56,723	56,723	56,723	56,723

Notes: Baseline sample of 56,723 defendant-case level observations filed 2010-2020. Columns (1) - (3) include controls for *court* × *year* × *subject* fixed effects. Column (4) includes controls for *court* + *year* + *subject* fixed effects. Standard errors are clustered at *court* level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.
Source: Prepared by the author.

Tabela 1.3 – Testing for random assignment of cases to courtrooms

	Pr(conviction) (1)	Courtroom stringency (2)
Age	-0.003***	5.15e-6
	(0.0004)	(5.42e-5)
Female	-0.072***	-0.005**
	(0.006)	(0.003)
Black race	0.017***	0.0003
	(0.006)	(0.0008)
Indigenous race	0.033	-0.012
	(0.053)	(0.012)
Non identified race	0.010	-0.0003
	(0.012)	(0.003)
<= 12 years education	-0.002	0.0002
	(0.013)	(0.002)
> 12 years education	-0.021*	0.002
	(0.011)	(0.002)
Missing Xs	-0.026***	-0.001
	(0.007)	(0.002)
Worked, t-1	-0.072***	-0.0008
	(0.009)	(0.002)
Worked, t-2 to t-3	-0.001	-0.0003
	(0.004)	(0.001)
Worked, t-4 to t-5	-0.015***	-0.0006
	(0.004)	(0.0009)
Court-Year-Subject FE	Yes	Yes
F (joint nullity), stat.	24.772	0.74355
F (joint nullity), p-value	7.29e-52	0.69717
Observations	56,723	56,723

Notes: Baseline sample of criminal cases processed 2010-2020. Standard errors are clustered at the *court* level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.
Source: Prepared by the author.

Tabela 1.4 – Test for monotonicity assumption

	Baseline instrument (1)	Reverse-sample instrument (2)
Sub-sample: conviction - 1st quartile		
Estimate	0.6808***	0.5473***
	(0.0422)	(0.0815)
Dependent mean	0.1236	0.1089
Observations	34,760	16,922
Sub-sample: conviction - 2nd quartile		
Estimate	0.9411***	0.8041***
	(0.0271)	(0.0431)
Dependent mean	0.4324	0.4300
Observations	34,761	16,921
Sub-sample: conviction - 3rd quartile		
Estimate	0.7711***	0.5652***
	(0.0386)	(0.0582)
Dependent mean	0.7602	0.7660
Observations	34,758	16,921
Sub-sample: conviction - 4th quartile		
Estimate	0.4437***	0.1924***
	(0.0649)	(0.0501)
Dependent mean	0.9307	0.9518
Observations	34,763	16,922
Sub-sample: previously non-employed		
Estimate	0.7904***	0.2076***
	(0.0339)	(0.0572)
Dependent mean	0.6032	0.4277
Observations	91,707	15,204
Sub-sample: previously employed		
Estimate	0.8151***	0.3768***
	(0.0292)	(0.0579)
Dependent mean	0.4813	0.4940
Observations	47,335	19,183
Sub-sample: age >= 35		
Estimate	0.8140***	0.4035***
	(0.0421)	(0.0495)
Dependent mean	0.3854	0.3926
Observations	42,402	16,827
Sub-sample: age < 35		
Estimate	0.7981***	0.4389***
	(0.0295)	(0.1013)
Dependent mean	0.6391	0.2729
Observations	96,640	10,313
Sub-sample: < 9 years of education		

(continued)

	Baseline instrument (1)	Reverse-sample instrument (2)
Estimate	0.9396***	0.2723*
	(0.1604)	(0.1424)
Dependent mean	0.4863	0.2960
Observations	3,037	581
Sub-sample: < 12 years of education		
Estimate	0.9348***	0.4765**
	(0.0665)	(0.1967)
Dependent mean	0.5428	0.2987
Observations	9,539	1,269
Sub-sample: >= 12 years of education		
Estimate	0.7940***	0.4987*
	(0.0284)	(0.0711)
Dependent mean	0.5650	0.1392
Observations	126,466	668
Sub-sample: black race		
Estimate	0.8375***	0.2911***
	(0.0396)	(0.0922)
Dependent mean	0.5938	0.5811
Observations	26,372	6,469
Sub-sample: non-black race		
Estimate	0.7985***	0.1623***
	(0.0272)	(0.0223)
Dependent mean	0.5542	0.6403
Observations	112,670	6,434

Notes: We estimate an OLS regression of the probability of conviction on all the variables listed in Table1.1 to create an index representing the predicted probability of conviction used in panels A-D. Each column estimates the first stage for the category indicated in the panel. The baseline instrument is constructed as the leave-one-out average of the conviction rate. The reverse-sample instrument is created excluding all cases within the sub-sample listed in the panel. All specifications include *court* × *year* × *subject* fixed effects. Standard errors are clustered at *court* level. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.
Source: Prepared by the author.

Tabela 1.5 – Estimates of conviction on labor outcomes

	Pr(Work) (1)	Pr(Ever Work) (2)	Total Days Worked (3)	Cum. Total Days Worked (4)	Total Earnings (5)	Cum. Total Earnings (6)
OLS (all controls)	-0.071***	-0.067***	-0.468***	-0.541***	-0.141***	-0.219***
	(0.004)	(0.006)	(0.029)	(0.035)	(0.012)	(0.016)
RF (all controls)	-0.069***	-0.101***	-0.547***	-0.658***	-0.112**	-0.218***
	(0.024)	(0.026)	(0.151)	(0.179)	(0.047)	(0.073)
IV (no controls)	-0.081**	-0.118***	-0.643***	-0.772***	-0.137*	-0.263**
	(0.035)	(0.038)	(0.232)	(0.272)	(0.073)	(0.115)
IV (all controls)	-0.081***	-0.119***	-0.645***	-0.776***	-0.133**	-0.258***
	(0.029)	(0.032)	(0.184)	(0.217)	(0.057)	(0.088)
Dependent mean	0.432	0.552	3.14	3.75	0.631	1.11
Observations	42,597	42,597	42,597	42,597	38,767	38,767

Notes: Baseline estimation sample of criminal cases filed 2010-2017. I.H.S stands for *Inverse Hyperbolic Sine*. Total Earnings and Cum. Total Earnings are expressed in units of thousands of *Reais*. All estimations include controls for *court x year x subject* fixed effects. Standard errors are clustered at *court* level. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Source: Prepared by the author.

Tabela 1.6 – Estimates of conviction on criminal recidivism outcomes

	Pr(Ever Charged)	I.H.S. Cum. Charges	Pr(Ever Charged) Non-severe	I.H.S. Cum. Charges Non-severe	Pr(Ever Charged) Severe	I.H.S. Cum. Charges Severe
	(1)	(2)	(3)	(4)	(5)	(6)
OLS (all controls)	0.166***	0.186***	0.097***	0.102***	0.042***	0.040***
	(0.008)	(0.009)	(0.008)	(0.008)	(0.006)	(0.006)
RF (all controls)	0.107***	0.111***	0.040*	0.041*	0.021	0.015
	(0.033)	(0.038)	(0.021)	(0.024)	(0.015)	(0.015)
IV (no controls)	0.126***	0.130***	0.052*	0.054*	0.028	0.021
	(0.040)	(0.046)	(0.027)	(0.031)	(0.021)	(0.020)
IV (all controls)	0.127***	0.131***	0.051*	0.053*	0.028	0.020
	(0.039)	(0.045)	(0.027)	(0.031)	(0.020)	(0.020)
Dependent mean	0.200	0.220	0.081	0.083	0.044	0.042
Observations	42,597	42,597	30,740	30,740	28,742	28,742

Notes: Baseline estimation sample of criminal cases filed 2010-2017. I.H.S stands for *Inverse Hyperbolic Sine*. All estimations include controls for *court x year x subject* fixed effects. Standard errors are clustered at *court* level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.
Source: Prepared by the author.

Tabela 1.7 – Heterogeneity estimation

		Crime Level		Employment		Gender		Age	
	All	Low severity	High severity	Previously unemployed	Previously employed	Male	Female	Under 35	Over 35
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Pr(Work)									
IV (all controls)	-0.081***	-0.058*	-0.185**	-0.117**	-0.085**	-0.081***	-0.162	-0.136***	-0.020
	(0.029)	(0.034)	(0.079)	(0.049)	(0.042)	(0.031)	(0.127)	(0.034)	(0.046)
Dependent mean	0.432	0.467	0.224	0.194	0.619	0.434	0.419	0.410	0.461
Observations	42,597	30,117	6,385	18,721	23,876	37,754	4,843	23,985	18,612
Panel B: Pr(Ever Worked)									
IV (all controls)	-0.119***	-0.105***	-0.190*	-0.151**	-0.132***	-0.121***	-0.136	-0.153***	-0.068
	(0.032)	(0.039)	(0.110)	(0.064)	(0.041)	(0.032)	(0.168)	(0.045)	(0.054)
Dependent mean	0.552	0.591	0.329	0.305	0.746	0.553	0.548	0.537	0.571
Observations	42,597	30,117	6,385	18,721	23,876	37,754	4,843	23,985	18,612
Panel C: Total Days Worked									
IV (all controls)	-0.645***	-0.555**	-1.17**	-0.799**	-0.772***	-0.651***	-0.933	-0.897***	-0.324
	(0.184)	(0.224)	(0.537)	(0.346)	(0.261)	(0.192)	(0.925)	(0.241)	(0.307)
Dependent mean	3.14	3.37	1.75	1.57	4.37	3.15	3.08	3.00	3.32
Observations	42,597	30,117	6,385	18,721	23,876	37,754	4,843	23,985	18,612
Panel D: I.H.S. Cum. Total Days Worked									
IV (all controls)	-0.776***	-0.671**	-1.38**	-0.964**	-0.917***	-0.783***	-1.08	-1.06***	-0.399
	(0.217)	(0.266)	(0.652)	(0.416)	(0.304)	(0.226)	(1.11)	(0.289)	(0.365)
Dependent mean	3.75	4.02	2.11	1.90	5.19	3.75	3.68	3.59	3.95
Observations	42,597	30,117	6,385	18,721	23,876	37,754	4,843	23,985	18,612
Panel E: I.H.S. Total Earnings (real)									
IV (all controls)	-0.133**	-0.096	-0.278***	-0.136**	-0.123	-0.164**	-0.071	-0.189***	-0.049
	(0.057)	(0.064)	(0.105)	(0.069)	(0.084)	(0.064)	(0.172)	(0.057)	(0.102)

(continued)

		Crime Level		Employment		Gender		Age	
	All	Low severity	High severity	Previously unemployed	Previously employed	Male	Female	Under 35	Over 35
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent mean	0.631	0.696	0.255	0.257	0.960	0.639	0.569	0.570	0.712
Observations	38,767	27,361	5,973	18,123	20,644	34,379	4,388	22,003	16,764
Panel F: I.H.S. Cum. Total Earnings (real)									
IV (all controls)	-0.258***	-0.205**	-0.465***	-0.271**	-0.253**	-0.301***	-0.178	-0.344***	-0.110
	(0.088)	(0.102)	(0.170)	(0.126)	(0.124)	(0.097)	(0.319)	(0.094)	(0.156)
Dependent mean	1.11	1.22	0.499	0.503	1.65	1.13	1.03	1.03	1.22
Observations	38,767	27,361	5,973	18,123	20,644	34,379	4,388	22,003	16,764
Panel G: Pr(Ever Charged)									
IV (all controls)	0.128***	0.154***	0.010	0.135**	0.127***	0.138***	-0.074	0.099	0.162***
	(0.033)	(0.047)	(0.105)	(0.058)	(0.045)	(0.043)	(0.093)	(0.061)	(0.040)
Dependent mean	0.207	0.196	0.235	0.238	0.170	0.207	0.141	0.222	0.171
Observations	52,894	30,117	6,385	18,721	23,876	37,754	4,843	23,985	18,612
Panel H: I.H.S. Cum. Number of New Charges									
IV (all controls)	0.138***	0.168***	-0.031	0.141**	0.124**	0.152***	-0.200*	0.099	0.164***
	(0.035)	(0.053)	(0.117)	(0.062)	(0.051)	(0.051)	(0.106)	(0.067)	(0.048)
Dependent mean	0.231	0.216	0.248	0.265	0.185	0.229	0.154	0.247	0.186
Observations	52,894	30,117	6,385	18,721	23,876	37,754	4,843	23,985	18,612
Panel I: Pr(Ever Non-Severe Charged)									
IV (all controls)	0.054**	0.068**	0.064	0.068	0.028	0.068**	-0.167*	0.059	0.060*
	(0.024)	(0.033)	(0.045)	(0.048)	(0.029)	(0.032)	(0.096)	(0.044)	(0.035)
Dependent mean	0.077	0.094	0.047	0.099	0.067	0.087	0.034	0.092	0.068
Observations	38,687	22,650	3,864	12,826	17,914	26,989	3,751	16,526	14,214

(continued)

	All	Crime Level		Employment		Gender		Age	
		Low severity	High severity	Previously unemployed	Previously employed	Male	Female	Under 35	Over 35
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel J: I.H.S. Cum. Number of New Non-Severe Charges									
IV (all controls)	0.054**	0.081**	0.053	0.056	0.037	0.077**	-0.292**	0.053	0.058
	(0.027)	(0.037)	(0.040)	(0.050)	(0.029)	(0.036)	(0.145)	(0.051)	(0.038)
Dependent mean	0.079	0.097	0.045	0.103	0.069	0.089	0.037	0.094	0.070
Observations	38,687	22,650	3,864	12,826	17,914	26,989	3,751	16,526	14,214
Panel L: Pr(Ever Severe Charged)									
IV (all controls)	0.024	0.052**	-0.155	-0.013	0.057**	0.031	-0.062*	0.058	-0.008
	(0.016)	(0.021)	(0.140)	(0.037)	(0.025)	(0.024)	(0.037)	(0.042)	(0.018)
Dependent mean	0.044	0.040	0.090	0.060	0.033	0.048	0.018	0.061	0.023
Observations	36,467	20,404	4,290	11,972	16,770	25,104	3,638	15,741	13,001
Panel M: I.H.S. Cum. Number of New Severe Charges									
IV (all controls)	0.019	0.045**	-0.170	-0.013	0.046*	0.023	-0.055	0.046	-0.009
	(0.016)	(0.020)	(0.150)	(0.035)	(0.026)	(0.024)	(0.035)	(0.038)	(0.016)
Dependent mean	0.042	0.038	0.084	0.057	0.031	0.045	0.017	0.058	0.022
Observations	36,467	20,404	4,290	11,972	16,770	25,104	3,638	15,741	13,001

Notes: Baseline estimation sample of criminal cases filed 2010-2017. I.H.S stands for *Inverse Hyperbolic Sine*. Total Earnings and Cum. Total Earnings are expressed in units of thousands of *Reais*. All estimations include controls for *court x year x subject* fixed effects. Standard errors are clustered at *court* level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Source: Prepared by the author.

Tabela 1.8 – Robustness Checks

	Baseline	>= 5 cases	>= 15 cases	>= 20 cases	>=25 cases	Split -sample
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Pr(conviction)						
First stage	0.827***	0.764***	0.847***	0.855***	0.856***	0.727***
	(0.029)	(0.027)	(0.032)	(0.038)	(0.042)	(0.042)
Dependant mean	0.483	0.478	0.484	0.486	0.489	0.491
Observations	56,723	72,874	45,787	37,826	31,924	19,524
Panel B. Pr(Work)						
RF (all controls)	-0.069***	-0.048**	-0.079***	-0.083***	-0.069*	-0.047
	(0.024)	(0.018)	(0.028)	(0.032)	(0.040)	(0.039)
IV (all controls)	-0.081***	-0.060**	-0.092***	-0.097**	-0.080*	-0.065
	(0.029)	(0.023)	(0.034)	(0.038)	(0.048)	(0.054)
Dependent mean	0.432	0.437	0.427	0.424	0.423	0.435
Observations	42,597	54,671	34,409	28,406	24,054	19,524
Panel C. Pr(Ever Worked)						
RF (all controls)	-0.101***	-0.065***	-0.103***	-0.108***	-0.093**	-0.059
	(0.026)	(0.022)	(0.029)	(0.035)	(0.043)	(0.049)
IV (all controls)	-0.119***	-0.082***	-0.119***	-0.126***	-0.108**	-0.081
	(0.032)	(0.028)	(0.035)	(0.043)	(0.051)	(0.067)
Dependent mean	0.552	0.559	0.546	0.542	0.540	0.554
Observations	42,597	54,671	34,409	28,406	24,054	19,524
Panel D. I.H.S. Total Days Worked						
RF (all controls)	-0.547***	-0.369***	-0.545***	-0.583***	-0.511*	-0.361
	(0.151)	(0.126)	(0.174)	(0.207)	(0.260)	(0.263)
IV (all controls)	-0.645***	-0.465***	-0.633***	-0.681***	-0.592*	-0.496
	(0.184)	(0.161)	(0.211)	(0.252)	(0.309)	(0.363)
Dependent mean	3.14	3.18	3.10	3.08	3.07	3.15
Observations	42,597	54,671	34,409	28,406	24,054	19,524
Panel E. I.H.S. Cum. Total Days Worked						
RF (all controls)	-0.658***	-0.441***	-0.657***	-0.700***	-0.612**	-0.425
	(0.179)	(0.149)	(0.204)	(0.244)	(0.306)	(0.315)
IV (all controls)	-0.776***	-0.556***	-0.764***	-0.818***	-0.709*	-0.584
	(0.217)	(0.191)	(0.248)	(0.298)	(0.364)	(0.435)
Dependent mean	3.75	3.79	3.70	3.67	3.66	3.76
Observations	42,597	54,671	34,409	28,406	24,054	19,524
Panel F. I.H.S. Total Earnings						
RF (all controls)	-0.112**	-0.060*	-0.102**	-0.129**	-0.102	-0.044
	(0.047)	(0.033)	(0.051)	(0.059)	(0.076)	(0.075)
IV (all controls)	-0.133**	-0.075*	-0.118*	-0.149**	-0.116	-0.060

(continued)

	Baseline (1)	>= 5 cases (2)	>= 15 cases (3)	>= 20 cases (4)	>=25 cases (5)	Split -sample (6)
	(0.057)	(0.041)	(0.062)	(0.072)	(0.089)	(0.102)
Dependent mean	0.631	0.639	0.623	0.620	0.618	0.637
Observations	38,767	49,728	31,371	25,894	21,914	17,792
Panel G. I.H.S. Cum. Total Earnings						
RF (all controls)	-0.218***	-0.125**	-0.199**	-0.237**	-0.191	-0.092
	(0.073)	(0.053)	(0.081)	(0.094)	(0.118)	(0.117)
IV (all controls)	-0.258***	-0.156**	-0.231**	-0.274**	-0.217	-0.125
	(0.088)	(0.066)	(0.099)	(0.115)	(0.139)	(0.160)
Dependent mean	1.11	1.13	1.10	1.09	1.09	1.12
Observations	38,767	49,728	31,371	25,894	21,914	17,792
Panel H. Pr(Ever reoffending)						
RF (all controls)	0.107***	0.133***	0.130***	0.122***	0.118**	0.119***
	(0.033)	(0.026)	(0.037)	(0.046)	(0.049)	(0.035)
IV (all controls)	0.127***	0.167***	0.151***	0.142**	0.137**	0.164***
	(0.039)	(0.032)	(0.043)	(0.055)	(0.056)	(0.048)
Dependent mean	0.200	0.201	0.198	0.197	0.195	0.195
Observations	42,597	54,671	34,409	28,406	24,054	19,524
Panel I. I.H.S. Cum. Number of Charges						
RF (all controls)	0.111***	0.145***	0.144***	0.136**	0.132**	0.128***
	(0.038)	(0.030)	(0.046)	(0.058)	(0.062)	(0.041)
IV (all controls)	0.131***	0.183***	0.168***	0.159**	0.152**	0.176***
	(0.045)	(0.038)	(0.053)	(0.069)	(0.071)	(0.056)
Dependent mean	0.220	0.224	0.219	0.218	0.216	0.214
Observations	42,597	54,671	34,409	28,406	24,054	19,524
Panel J. Pr(Ever reoffending - Non-severe cases)						
RF (all controls)	0.040*	0.042**	0.064***	0.060**	0.050*	0.046*
	(0.021)	(0.018)	(0.022)	(0.024)	(0.028)	(0.027)
IV (all controls)	0.051*	0.058**	0.081***	0.079**	0.065*	0.064*
	(0.027)	(0.024)	(0.029)	(0.032)	(0.038)	(0.038)
Dependent mean	0.081	0.081	0.079	0.076	0.073	0.078
Observations	30,740	39,251	25,003	20,614	17,463	14,100
Panel L. I.H.S. Cum. Number of Non-Severe Charges						
RF (all controls)	0.041*	0.044**	0.061**	0.058*	0.043	0.037
	(0.024)	(0.020)	(0.029)	(0.031)	(0.034)	(0.031)
IV (all controls)	0.053*	0.061**	0.078**	0.076*	0.057	0.052
	(0.031)	(0.028)	(0.037)	(0.041)	(0.045)	(0.044)
Dependent mean	0.083	0.083	0.082	0.078	0.075	0.079
Observations	30,740	39,251	25,003	20,614	17,463	14,100

(continued)

	Baseline	>= 5 cases	>= 15 cases	>= 20 cases	>=25 cases	Split -sample
	(1)	(2)	(3)	(4)	(5)	(6)
Panel M. Pr(Ever reoffending - Severe cases)						
RF (all controls)	0.021	0.016	0.015	0.015	0.015	0.042
	(0.015)	(0.015)	(0.018)	(0.023)	(0.024)	(0.029)
IV (all controls)	0.028	0.023	0.019	0.021	0.020	0.063
	(0.020)	(0.021)	(0.023)	(0.031)	(0.031)	(0.043)
Dependent mean	0.044	0.042	0.045	0.046	0.046	0.047
Observations	28,742	36,450	23,448	19,460	16,526	13,252
Panel N. I.H.S. Cum. Number of Severe Charges						
RF (all controls)	0.015	0.012	0.009	0.009	0.011	0.040
	(0.015)	(0.014)	(0.018)	(0.023)	(0.023)	(0.027)
IV (all controls)	0.020	0.018	0.012	0.012	0.015	0.060
	(0.020)	(0.021)	(0.023)	(0.031)	(0.030)	(0.041)
Dependent mean	0.042	0.040	0.042	0.044	0.043	0.044
Observations	28,742	36,450	23,448	19,460	16,526	13,252

Notes: Column (1) shows baseline estimates using leave-out mean courtroom stringency including cases assigned to the courtroom that have handled at least 10 cases of a subject within a year. In columns (2)-(5), courtrooms are required to handle at least 5, 10, 15, 20, and 25 cases of a subject within a year, respectively. Column (6) employs a three-step process to estimate the IV model outlined in equations 1.2-1.3. Firstly, the baseline estimation sample is randomly divided into two mutually exclusive sub-samples. Secondly, one of these sub-samples is selected and the instrument is constructed using each judge's case decisions in the other sub-sample. Finally, the retained sub-sample is utilized to estimate the IV model. I.H.S stands for *Inverse Hyperbolic Sine*. Total Earnings and Cum. Total Earnings are expressed in units of thousands of *Reais*. All estimations include all controls in Table 1.1 and *court x year x subject* fixed effects. Standard errors are clustered at *court* level.*p<0.1, **p<0.05, ***p<0.01.
Source: Prepared by the author.

2 CONVICTION'S ECHO: UNVEILING FAMILY CONSEQUENCES IN BRAZIL

2.1 INTRODUCTION

The experience of an individual's conviction and subsequent incarceration is a deeply challenging and transformative event that reverberates throughout the lives of family members. Capturing these effects is particularly relevant from a policy perspective as the societal costs and benefits of incarceration could be magnified or muted once the effect on the families is taken into account. While the literature on this subject has flourished in recent years, it remains a complex and nuanced field, as there are arguments for both positive (BHULLER *et al.*, 2018a; ARTEAGA, 2021; NORRIS; PECENCO; WEAVER, 2021) and negative effects (DOBBIE *et al.*, 2019) of an individual's conviction on family members. Many studies have examined the effect of imprisonment on a broad range of topics, such as employment, earnings, education, health, and many others, in the US and Europe. Despite the high incarceration rate in Latin America, evidence about the causal impact on convicts and their families in the region is scarce.

Estimating the causal impact of a criminal conviction on family dynamics poses some challenges for researchers. Firstly, the lack of individual-level panel data on criminal records and the difficulty of linking individuals to their respective family units restricts the ability to examine how a criminal conviction directly affects different family members. Additionally, the issue of endogeneity poses complexities in establishing a causal relationship between a criminal conviction and family outcomes. Individuals from the family of convicted defendants may possess unobservable characteristics that differentiate them from the non-convicted counterparts, confounding the true effect of the criminal conviction on family dynamics. Addressing these challenges requires the implementation of rigorous quasi-experimental designs and identification strategies to estimate the causal impact of an individual's criminal conviction on subsequent family outcomes.

In this paper, we tackle these data and methodological challenges within the context of Brazil and examine the causal effect of individuals' convictions on the labor, criminal, and education outcomes of their family members. We address the data challenge by leveraging a comprehensive linkage of administrative databases in Brazil. We begin by merging our collected criminal case database with the Brazilian Registry for Social Programs (*Cadastro Único*) to establish the connection between defendants and their family units. We then link these records to the universe of formal workers (*RAIS*) and the school census (*Censo Escolar*) to construct a unique panel dataset that allows us to track labor, criminal and education outcomes for family members of individuals involved in the criminal justice.

To address the endogeneity issue, we leverage the institutional rule in Brazil that dictates the random assignment of judicial cases to courtrooms, which exhibit systematic variations in their tendency to convict individuals. We construct the courtroom stringency measure as the leave-one-out average of the criminal conviction rate and use it as an instrument for conviction decisions. By applying this instrumental variable design, we can estimate the local average treatment effect of individuals' convictions on labor, criminal, and education outcomes of their family members.

Our findings provide valuable insights into the far-reaching consequences of criminal convictions and shed light on the potential heterogeneity in these effects across different subgroups. First, the results revealed the adverse effects of individuals' criminal convictions on the labor market prospects of family members. We estimate that conviction decreases the probability of employment (extensive margin) by 9 percentage points after two years from the start of the case. This effect is considerable, amounting to a 20% drop from the average baseline employment level of 0.46. Moreover, we observe a persistent and noteworthy reduction of over 10% in earnings within two years following a case filing. These results highlight the substantial negative impact of individual's criminal convictions on the employment prospects of their family members. Additionally, our analysis identifies that male individuals within the family are particularly vulnerable to these adverse labor market effects.

Second, our results also reveal a significant increase in criminal activity among family members following an individuals' conviction, with a particular emphasis on severe crimes. We estimate that the likelihood of a family member being charged for a crime increases by 3.6 percentage points. This result is notable considering the relatively low baseline rate of criminal behavior among family members in our sample, effectively doubling the chances of engaging in criminal activities for those individuals whose relatives were marginally convicted. Moreover, we further examine criminal behavior by differentiating between the severity of cases. Our results demonstrate a distinct pattern, with a 3.8 percentage point increase in criminal activity specifically related to severe cases, while the estimate for non-severe cases remains close to zero. This suggests that the impact of individuals' criminal convictions primarily affects the occurrence of severe offenses within their family network. Our heterogeneity analysis found that siblings were particularly susceptible to engaging in more serious criminal activities, highlighting the influence of familial ties in transmitting criminal tendencies.

Finally, our analysis explored the effects of parental convictions on education outcomes. Our examination indicates a tendency toward lower education performance among family members affected by conviction. However, the results did not reach statistical significance, possibly due to the reduced sample size. Nevertheless, the overarching results suggest a potential negative influence of convictions on education outcomes.

Our research is closely linked to several studies that utilize quasi-random judge assignment to examine the family spillover of incarceration on multiple outcomes in different contexts. For instance, (BHULLER *et al.*, 2018b) finds that a father's incarceration has no effect on child's criminal activity or performance in school, while (BHULLER *et al.*, 2018a) finds a positive spillover effect of incarceration on the likelihood of siblings' criminal activity in Norway. In Sweden, (DOBBIE *et al.*, 2019) estimate that parental incarceration leads to increases in criminal activity, pregnancy, and worse employment. In the context of Finland, (HUTTUNEN *et al.*, 2019) estimate the impact of three types of punishments of parents (fines, probation, and prison) on child's outcomes in Finland and find mixed results. (NORRIS; PECENCO; WEAVER, 2021) and (ARTEAGA, 2021) estimate the effect of parental incarceration on children and found beneficial effects on some children outcomes in US and Colombia, respectively. Our study contributes significantly to this line of research as the majority of previous studies have used data from the U.S and Nordic countries. In contrast, our paper provides the first set of causal evidence of the effects of individuals' criminal convictions on their family dynamics in Brazil, the largest country in Latin America with the third-largest prisoner population[1] in the world.

The structure of our paper is as follows. Section 2.2 outlines our research design. We describe our data and sample selection process in Section 2.3. Our main results for labor, criminal, and education outcomes, as well as our heterogeneity, are presented in Section 2.4. Finally, we conclude in Section 2.5.

2.2 RESEARCH DESIGN

We aim to estimate the impact of an individual's criminal conviction on labor, criminal behavior, and education outcomes of their family members. We can represent the model relating future outcomes for family members to the indicator of individual's conviction as follows:

[1] World Prison Brief (2021) available at www.prisonstudies.org.

$$Y_{j,i,t} = \beta_t \mathrm{I}_i + X'\gamma + e_{j,i,t}, \tag{2.1}$$

where i represents individuals facing prosecution, j is a member of his family, t is the time of observation, β_t is the causal effect of interest, I_i is an indicator equal to 1 if the defendant i is convicted, X is a vector including controls variables for both defendants and their family members, $Y_{j,i,t}$ is the outcome of interest of individual j from the defendant's family i, measured t periods after case starts and $e_{j,i,t}$ is the error term. The problem of estimating Equation 2.1 is that any causal interpretation of β_t will be biased if conviction status is somehow correlated to any unobservable determinant of Y.

To address the issue of endogeneity, we employ a strategy that takes advantage of the random assignment of criminal cases to different courtrooms in Brazil, which exhibit systematic differences in their propensity to convict. This random variation in the probability of conviction, determined by the courtroom assigned to a defendant's case, allows us to identify the causal impact of parental conviction on family members. More details on the institutional setting in Brazil can be found in our prior work (FARIAS; SAMPAIO; BRITTO, 2023).

Formally, we identify the causal impact of a parental conviction on family members β_t using a measure of courtroom stringency (z) as an instrumental variable for being convicted. Our main specification is based on two-stage least squares (2SLS) estimation of β_t with the following two-equations system:

$$I_i = \delta z_{c,i} + X'\theta + \varepsilon_i, \tag{2.2}$$

$$Y_{j,i,t} = \beta_t I_i + X'\gamma + e_{j,i,t}, \tag{2.3}$$

where $z_{c,i}$ is our measure of leniency of the courtroom c assigned to defendant i's case and X is a vector, including a full set of court-subject-year fixed effects representing the level at which randomization of courtrooms occurs, as well as a set of control variables for both defendants and their family members.

Following standard practices in research on judge fixed effects, we construct our instrument as the leave-one-out mean conviction rate. This instrumental variable averages the conviction rates in other cases handled by the courtroom, excluding the specific case under study. This approach helps to eliminate any correlation between the courtroom's ruling in a particular case and the value of the instrument. For each defendant i, we calculate a measure of stringency based on the initial courtroom assigned to their case and use it as an instrument for the conviction status.

Our research design is the same used in our prior work (FARIAS; SAMPAIO; BRITTO, 2023) and is consistent with previous studies such as (JR, 2008; TELLA; SCHARGRODSKY, 2013; MAESTAS; MULLEN; STRAND, 2013; DAHL; KOSTØL; MOGSTAD, 2014; FRENCH; SONG, 2014; AIZER; JR, 2015; DOBBIE; SONG, 2015; DOBBIE; GOLDIN; YANG, 2018; COHEN; YANG, 2019; HUTTUNEN *et al.*, 2019; DOBBIE *et al.*, 2019; HUTTUNEN; KAILA; NIX, 2020; BHULLER *et al.*, 2020; ARTEAGA, 2021; BHULLER; KHOURY; LøKEN, 2021; NORRIS; PECENCO; WEAVER, 2021; COLLINSON *et al.*, 2022), which have investigated related topics using similar approaches.

Assuming the exogeneity and monotonicity assumptions hold for our instrumental variable, we can interpret the parameter β_t in Equation 2.3 as the local average treatment effect (LATE) of parental conviction on family members. The LATE represents the causal impact of an individual's criminal conviction for family members whose defendants would have received a different decision if their case had been assigned to a different courtroom.

2.3 DATA

In order to estimate the impact of an individual's conviction on labor, criminal behavior, and education among family members, we performed a unique merge between criminal cases and a rich set of administrative data in Brazil.

Data on criminal cases in Brazil is the same used in (FARIAS; SAMPAIO; BRITTO, 2023). The criminal data is comprised of two sources. The primary source was text sentences from all criminal adjudicated cases filed at the State Court of São Paulo (*Tribunal de Justiça de São Paulo, TJSP*). The second source of information was proprietary data from a private firm that collects judicial data from multiple Brazilian courts. By combining these two datasets, we have a comprehensive picture of criminal cases in Brazil that allows us both to measure the treatment variable (convicted or not) as well as to track future criminal behavior. Further details on how final decisions are extracted and how we enhance the criminal case with unique identifiers (*CPF*) for the defendants can be found in (FARIAS; SAMPAIO; BRITTO, 2023). Overall, we are able to retrieve 2,814,081 criminal case decisions and 928,157 uniquely identified defendants from 2010-2022 period.

To establish the link between defendants and their family members, we use the

Brazilian Registry for Social Programs, referred to as *Cadastro Único*. This dataset encompasses individuals participating in federal social programs throughout the country. Recognized as the primary instrument for administering several social programs such as Bolsa Família and many others, the *Cadastro Único* serves as the Brazilian census specifically targeting the low-income population. By utilizing this extensive registry, we gained valuable insights into the interconnectedness of defendants and their respective family members. Our analysis leveraged yearly snapshots of the data spanning from 2012 to 2019. With a coverage of approximately 75 million individuals and 25 million families as of 2019, the dataset provided a comprehensive representation of the population under scrutiny.

To analyze labor market outcomes in Brazil, our study relies on the *Relação Anual de Informações Sociais (RAIS)* spanning from 2002 to 2020. This dataset encompasses information on all formal workers and firms in the country. The RAIS dataset contains a wide range of valuable information, including job start and end dates, job locations, unique identifiers for employers and employees, contract types, occupation and sectoral codes, worker education, race, and earnings. By leveraging this dataset, we are able to construct meaningful measures of labor outcomes. Our analysis incorporates measures such as the total number of days worked in a year and total earnings, allowing us to gain deep insights into labor market dynamics in Brazil.

Finally, our study incorporates data from the school census (*Censo Escolar*), covering the period from 2007 to 2017, to track education outcomes among family members. This dataset offers comprehensive and detailed information on students, including enrollment status, grade levels, demographic characteristics, and school-specific attributes.

2.3.1 Linking Defendants to Family Members

Our study is faced with the challenge of linking defendants' family members to various data sources. To tackle this issue, we implemented a rigorous and systematic approach.

To establish the connection between defendants and their families, we utilized a two-step procedure that involved linking the defendants' dataset with the *Cadastro Único* using the *CPF* as a common identifier.

In the first step, we identified the family unit associated with each defendant in the *Cadastro Único* by retrieving the unique family identifier assigned to them. This allowed us to identify all the families to which the defendant belongs.

Subsequently, utilizing the family identifier obtained in the previous step, we retrieved all the family members associated with each defendant. Specifically, we focused on families where the defendant served as the head or families in which the defendant was identified as an offspring.

Our study successfully establishes accurate linkages between defendants and their family members, resulting in a robust dataset. Out of a total of 928,157 defendants, we are able to establish links for 128,854 defendants to their respective family members. This process enables us to construct a dataset that encompasses 208,639 family members and facilitates their connection to other relevant data sources such as RAIS and the criminal dataset.

By integrating these diverse data sources, our analysis gains depth and breadth, allowing us to examine the potential impact on family dynamics on the labor market and involvement in the criminal justice system.

2.3.2 Linking Family Members to *Censo Escolar*

To establish a connection between family members and the *Censo Escolar* dataset, we encounter a significant challenge due to the absence of a unique identifier like the CPF (Brazilian individual taxpayer registry number) in the latter registry. However, we can overcome this limitation by following the methodology proposed by (OLIVEIRA; SOARES, 2013). We proceed as follows.

First, using the *Censo Escolar* dataset, we generate a key variable called the "*INEP key*"for all students. This key is created by combining information such as the municipality of birth, school identifier, birth date, and gender. The *INEP key* serves as a unique identifier within the *Censo Escolar* dataset and enables us to track students over time.

Similarly, we are able to create the same *INEP key* for all family members using the information available in the *Cadastro Único* registry. Although the *INEP key* cannot be used to directly link with other datasets, it allows us to establish a common identifier between family members and students in *Censo Escolar*.

By following this procedure, we can associate the unique identifier (CPF) from *Cadastro Único* with each student for whom we were able to successfully merge the family member data with the *Censo Escolar* dataset. This allows us to establish the direct link between family members of defendants and their educational information.

As a result of this procedure, we have successfully linked 58,990 out of 208,639 family members to the *Censo Escolar* dataset, covering the period from 2007 to 2017. This linkage provides us with a valuable opportunity to examine the potential effects of an individual's conviction on the educational outcomes of family members.

2.3.3 Sample Selection

The dataset on criminal cases used in this study is the same used in (FARIAS; SAMPAIO; BRITTO, 2023). It consists of 2,814,081 criminal cases where a sentence was issued between 2010 and 2022. To refine the dataset, non-randomly assigned cases were excluded, and only courtrooms with at least 10 cases per year and subject matter during the specified period were included in the sample. This resulted in a sample of 579,684 randomly assigned cases from courtrooms that met the criteria. This sample is the same used in (FARIAS; SAMPAIO; BRITTO, 2023) to construct the instrument variable for the study.

To refine our estimation sample, we apply additional restrictions based on the specific outcome of interest. For estimating labor and criminal behavior outcomes, we focus on family members who were between 18 and 55 years old at the start of their relative's cases. This age range allows us to capture individuals who are typically in the working-age range and are more likely to have labor and criminal outcomes that can be meaningfully assessed. Furthermore, we limit our analysis to criminal cases that started between 2010 and 2017 period. This duration ensures that each family member can be tracked and followed for three years, providing a more comprehensive understanding of the potential effects of an individual's conviction on labor and criminal behavior outcomes.

For studying education outcomes, we narrow down the sample to family members aged from 6 years old at the start of the case. Additionally, we only include cases that started from the 2010 to 2014 period. By focusing on this age range and timeframe, we can specifically examine the impact of conviction on the educational outcomes of family members during their crucial schooling years.

After implementing the specified restrictions and merging with the instrument information, we obtain a baseline estimation sample for labor and criminal outcomes consisting of 9,847 cases, involving 9,936 defendants, across 805 courtrooms, and 13,191 family members. Similarly, for the estimation of educational outcomes, we have a sample of 1,705 cases, involving 1,722 defendants, across 471 courtrooms, and 2,182 family members.

2.3.4 Descriptive Statistics

Table 2.1 provides a description of the demographic characteristics of individuals connected to defendants involved in the criminal justice system during the 2010-2022 period.

[Tabela 2.1 about here.]

In Column 1, we examine the characteristics of all family members in our sample. The results indicate that the majority of individuals are siblings, comprising over 60% of the total. Children constitute approximately 25%, while spouses make up 12% of the sample. Gender distribution is relatively balanced, with males accounting for around 55% and females 45%. On average, these individuals are 21 years old at the time the cases are filed. The sample is predominantly composed of White individuals with at least a high school education (12 years or more).

Columns 2 and 3 further explore the characteristics of individuals linked to convicted defendants compared to those linked to non-convicted defendants. The analysis reveals a relatively balanced distribution across various variables, including gender, age, race, and education, indicating similarities between the two groups of family members. However, there are notable differences in the composition of family members. Specifically, family members linked to convicted defendants are more likely to be siblings (71% versus 47% for the non-convicted group) and less likely to be children (18% versus 36%) or spouses (9% versus 15%) compared to family members of non-convicted defendants.

These descriptive findings provide important insights into the demographic composition of family members connected to defendants within the Brazilian criminal justice system.

2.4 MAIN RESULTS

2.4.1 Instrument Validity

Our instrumental variable is constructed based on the average conviction rate in other cases handled by the courtroom, including both past and future cases outside of our estimation sample. The instrument exhibits a wide range of courtroom stringency rates, with values ranging from 0.32 to 0.62 across courtrooms at the 1^{st} and 99^{th} percentiles, respectively. The mean stringency rate is 0.47, with a standard deviation of 0.05. We provide a visual representation of this distribution in (FARIAS; SAMPAIO; BRITTO, 2023).

To assess the strength of our instrument, Table 2.2 presents the results of the first stage equation for both the sample of defendants linked to their family members and the overall defendant sample. The findings demonstrate a robust and highly significant relationship between the instrument and the conviction status. In both cases, a 10 percentage point higher probability of conviction in the assigned courtroom leads to an 8 percentage point increase in the likelihood of being convicted. This effect remains consistent across different samples, suggesting the instrument's strength and validity. Considering the baseline conviction rate of 0.6, this represents a substantial 13% deviation from the mean.

[Tabela 2.2 about here.]

In Table 2.3, we provide evidence that cases in our sample of defendants linked to their family members are randomly assigned to courtrooms. Column 1 presents the results of the regression of parental conviction on a set of control variables measured before the start of the case for defendants in which we were able to link with family members. In column 3, we reproduce the same results from our previous work (FARIAS; SAMPAIO; BRITTO, 2023) for all defendants. Columns 2 and 4 show a regression of courtroom stringency on the same set of characteristics in both samples. The results reveal that while the defendant's characteristics are highly predictive of the criminal conviction indicator, they do not have any noticeable effect on courtroom stringency. The results are similar between the two samples. The estimates indicate that the coefficients are close to zero, and the number of significant coefficients is not higher than what would be expected by chance. Furthermore, the coefficients do not exhibit joint significance, providing compelling evidence for conditional randomization. These findings strongly support the notion of random assignment in our study, bolstering the validity of our research design.

[Tabela 2.3 about here.]

In (FARIAS; SAMPAIO; BRITTO, 2023), we provide additional tests for monotonicity and find strong support for the validity of our instrument.

2.4.2 The Effect on Labor Outcomes

We initiate our study on the impact of an individual's criminal conviction on family members by examining its effect on labor outcomes. Specifically, we analyze the employment outcomes from both extensive and intensive perspectives. To capture the extensive margin, we utilize a binary variable that indicates whether a family member has ever been employed by the end of a given time period. In terms of the intensive margin, we measure the yearly (real) earnings, cumulative (real) earnings (expressed in units of thousands), and the cumulative number of days worked within a specified period. To facilitate interpretation, we follow (NORRIS; PECENCO; WEAVER, 2021) and transform the intensive margin outcomes using the inverse hyperbolic sine (I.H.S) function, allowing for the interpretation of the results as percent changes.

Figure 2.1 examines how conviction affects the defendant's family labor outcomes. Each line plots the coefficients from a within-period version of our 2SLS equations. In Table 4, we summarize these results while adding more elements to our analysis.

[Figura 2.1 about here.]

Figure 2.1a presents the estimates of the effects of an individual's conviction on the extensive margin of employment for family members. The findings reveal a significant decrease in the probability of formal employment among family members in Brazil in the years following the initiation of the case. For instance, Panel A of Table 2.4 shows a substantial 9-percentage point decline in employment probability within two years after the case starts. This effect is considerable, amounting to a 20% reduction from the average baseline employment level of 0.46.

Moreover, the analysis of the intensive margin of employment, as depicted in Figure 2.1b, provides further insight into the labor outcomes of family members. While the results do not reach statistical significance, there is a noticeable declining trend in the cumulative number of days worked for individuals from convicted defendants' families. This suggests that even for those who are able to find employment, there may be challenges in sustaining consistent and long-term work opportunities.

Examining earnings, Figures 2.1c and 2.1d provide further insights into the financial consequences of convictions. Panels C and D of Table 2.4 demonstrate a persistent and substantial reduction of over 10% in earnings within two years following a case filing. These findings underscore the enduring impact of individuals' criminal convictions on earning potential, leading to economic hardship for the family.

[Tabela 2.4 about here.]

Overall, our findings reveal a significant negative impact on labor outcomes, indicating that convictions have far-reaching consequences for the employment prospects of family members.

2.4.3 The Effect on Criminal Behavior

The previous analysis reveals that family members of individuals who are marginally convicted are likely to face negative consequences in terms of their labor market outcomes. However, an equally important concern is whether these family members are influenced by the criminal behavior of their convicted relative and more likely to engage in criminal activity themselves.

In this section, we investigate the extent to which family members of convicted individuals are prone to involvement in crime. We measure criminal activity by examining the number of criminal charges involving family members. Similar to our previous analyses, we consider both the extensive margin, represented by a binary variable indicating whether a family member has ever been criminally charged, and the intensive margin, taking the I.H.S of the cumulative number of criminal charges. Furthermore, we conduct separate estimations based on the severity[2] of the case to assess the differential effects.

The IV estimates depicted in Figure 2.2 and Table 2.5 provide valuable insights into the dynamics of criminal behavior among family members in response to relative's criminal convictions. These results shed light on both the extensive and intensive margins of criminal activity and highlight the differential effects based on the severity of the cases involved.

[Figura 2.2 about here.]

The results, as depicted in Figure 2.2a and Figure 2.2b, reveal a concerning trend of increasing criminal activity among family members over time. This trend is evident when examining both the extensive and intensive margins of criminal behavior. Notably, the figures underscore that the rise in criminal activity is predominantly observed in relation to severe crimes.

Panel A of Table 2.5 shows that within three years after the case starts, the probability of being charged for a crime increases 3.6 percentage points. Given the fact that only a small proportion of family members committed crime in our sample, this result doubles the chances of engaging in criminal activities for those individuals whose relative was marginally convicted. In Panels B and C, we further investigate criminal behavior focused on the severity of cases. The

[2] We adopt the same severity classification employed in our previous work (FARIAS; SAMPAIO; BRITTO, 2023) to categorize cases as either *severe* or *non-severe*. Cases with a base-penalty exceeding 4 years of sentence are classified as *severe* because they do not qualify for alternative sentencing options such as fines, community service, curfews, or other non-incarceration penalties. On the other hand, *non-severe* subjects are the ones with a base-penalty of less than 4 years of sentence, and are usually exchanged with non-incarceration options.

findings indicate an increase of 3.8 percentage points in criminal activity specifically related to severe cases, while the estimate for non-severe cases is close to zero.

Regarding the intensive margin of criminal behavior (Panels D, E, and F of Table 2.5), we observe a similar pattern, albeit with smaller effect sizes. This finding aligns with the notion that the impact of convictions is more pronounced for family members who are on the margin of committing a crime, rather than for those who are already serial offenders.

[Tabela 2.5 about here.]

Overall, our findings underscore the concerning link between individual's convictions and an elevated likelihood of criminal behavior among family members. The implications are significant, emphasizing the potential for the transmission of criminal tendencies within families. Moreover, these results raise potential concerns regarding the consequences of convictions, as it suggests that such convictions not only lead to an overall increase in criminal behavior among family members but also an inclination towards committing more serious offenses.

2.4.4 The Effect on Education Attainment

The detrimental effects of an individual's conviction can be amplified for the young members of the family, as they may face additional challenges. One particular area of interest is understanding how an individual's criminal conviction can impact the educational attainment of their children and siblings. On one hand, the combination of conviction and the resulting financial difficulties in the family can create an unstable environment that potentially has negative consequences for the family unit, especially the children. On the other hand, there are plausible reasons and results that point to the benefits of parental incarceration on them (ARTEAGA, 2021; NORRIS; PECENCO; WEAVER, 2021).

To examine the impact of criminal convictions on education outcomes, we delve into the relationship between an individual's conviction and the educational attainment of the family members. In our analysis, we focus on a specific subset of individuals: siblings and children who were at least 6 years old when the case began. Figure 2.3 and Table 2.6 depict the IV estimates.

[Figura 2.3 about here.]

Figure 2.3a presents the IV estimates of the impact of convictions on school enrollment. The results indicate a negative effect of convictions on the probability of children being enrolled in school. Panel A of Table 2.6 further quantifies this effect, showing a decrease of approximately 14 percentage points in school enrollment within three years after the case begins. This translates to a 17% decline from the baseline enrollment rate of 0.8. However, it's important to note that these estimates are not statistically significant, possibly due to the smaller sample size used in the analysis.

Similarly, Figure 2.3b illustrates the results for years of schooling. The analysis reveals a negative effect of individual's convictions on the number of years of schooling. Panel B of Table 2.6 indicates 0.7 years of schooling reduction within three years after the case starts. This result amounts to a 10% decrease from the average. Again, these estimates do not reach statistical significance.

Next, we investigate the probability of grade retention. Figure 2.3c indicates a positive trend on the chances of their children and siblings repeating a grade. In Table 2.6, Panel C shows that conviction induces 29.5, 22.6 and 48 percentage points after 1, 2 and 3 years the case starts, respectively. However, the estimates are not statistically significant, possibly because of the reduced sample size.

Lastly, we explore the impact of parental incarceration on age-grade distortion. We define age-grade distortion as when a student is behind his age group by two or more grades. Figure 2.3d depicts the results and shows the emergence of a positive trend, although the estimates are not statistically significant. Panel D of Table 2.6 displays a 23 percentage point increase within 3 years after filing, meaning conviction deepens age-grade distortion. Again, the estimates are not statistically significant.

[Tabela 2.6 about here.]

Overall, while our results suggest a negative association between an individual's criminal convictions and various educational outcomes, such as school enrollment, years of schooling, grade retention, and age-grade distortion, the statistical significance of these effects is limited by the reduced sample size. Although the estimated effects are not statistically significant, these results highlight the potential challenges faced by children and siblings of convicted individuals in their educational journeys. These results are in line with (DOBBIE *et al.*, 2019), but contrast with what was found in (ARTEAGA, 2021) and (NORRIS; PECENCO; WEAVER, 2021), which encountered beneficial effects on children's educational attainment.

2.4.5 Heterogeneity

In our main analysis, we estimate the local average treatment effect of individuals' criminal convictions on the labor, criminal, and education outcomes of their family members. The results revealed adverse effects on labor market outcomes, an increased likelihood of engaging in criminal activities, and potential negative impacts on education outcomes for individuals whose relative was marginally convicted.

To gain a deeper understanding of the heterogeneity of these effects, we conducted subgroup analyses based on various factors such as the relationship between the defendant and family member, gender, and the severity of the offense. These additional estimations provided valuable insights into how different subgroups within the sample were affected by a conviction, shedding light on potential variations in the magnitude of the effects.

Table 2.7 shows the IV results for the labor outcomes. We focused on the outcomes measured within 1-3 years after the start of the case. In Column 1, we replicate the results from our main analysis for comparison. Subsequent columns (2-5) explore how the effects vary across different subgroups within the family structure. While we observe larger effects for siblings and spouses, these results do not reach statistical significance. When considering gender differences (columns 6-7), we find that convictions have a more pronounced negative impact on the labor outcomes of male individuals within the family. Furthermore, by examining the type of crime committed by the convicted relative, we find that the effects are primarily concentrated among individuals whose parent committed a more serious offense (columns 8-9), but again these results are not statically significant.

[Tabela 2.7 about here.]

In Table 2.8, we show the IV results for the criminal behavior. Column 1, we replicate our results in which we found significant results for the extensive (criminally charge and being charged for severe crime) and intensive (number of severe crime charges) margins of crime. As we can observe from the table, the results for criminal behavior are concentrated among siblings, particularly towards in engaging in more serious offenses.

[Tabela 2.8 about here.]

Lastly, we focused on understanding how a parent's conviction can impact the educational attainment of their children and siblings. Table 2.9 presents the IV results for education outcomes. However, we did not find statistically significant results for any of the subgroups in terms of education outcomes. This lack of significance may be attributed to the reduced sample size, limiting the power to detect meaningful effects. Further research with larger sample sizes may provide additional insights into the potential effects of parental conviction on education outcomes for different subgroups.

[Tabela 2.9 about here.]

Overall, these heterogeneity analyses provide additional insights into how the effects of a conviction vary across different subgroups within the family. While some patterns and trends emerge, it is important to interpret these findings cautiously due to the lack of statistical significance in some cases. These insights can inform targeted interventions and policy measures that address the challenges faced by different individuals and their families in the aftermath of a conviction.

2.5 CONCLUSION

In this paper, we examined the impact of an individual's convictions on various outcomes for family members, including labor market outcomes, criminal behavior, and education attainment. Our findings provide valuable insights into the far-reaching consequences of convictions and shed light on the potential heterogeneity in these effects across different subgroups.

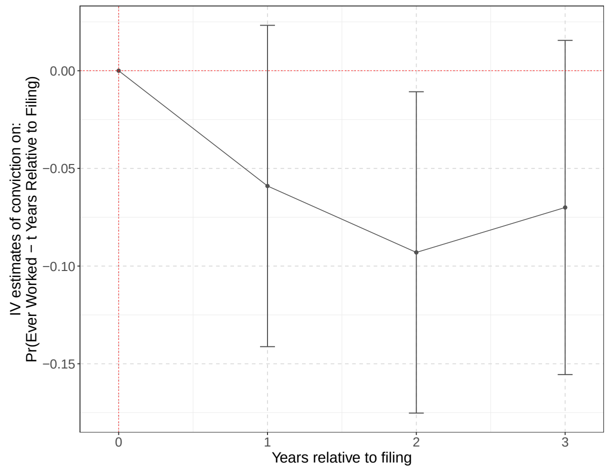
First, our main analysis revealed the adverse effects of an individual's criminal convictions on labor market prospects for their family members. Individuals with marginally convicted relatives experienced negative impacts on their employment status and earnings. Additionally, we found that male individuals within the family were more vulnerable.

Second, we examined the impact of criminal convictions on criminal behavior among family members. Our results indicated a significant increase in criminal activity following a conviction, particularly in terms of severe crimes. Siblings were particularly susceptible to engaging in more serious criminal activities, highlighting the influence of familial ties in transmitting criminal tendencies.

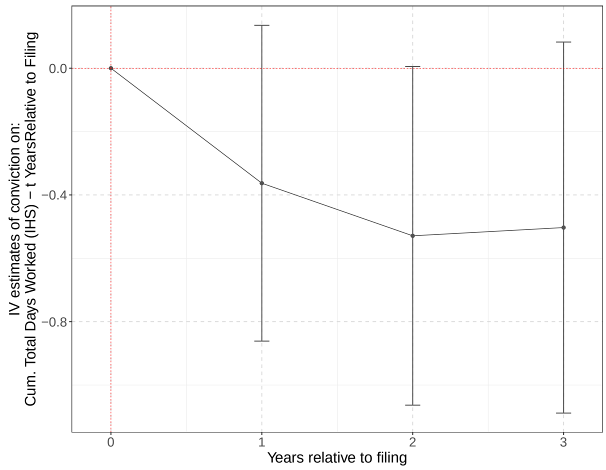
Finally, our analysis explored the effects of an individual's convictions on education outcomes for children and siblings. Our examination indicates a tendency toward lower education performance among family members affected by a conviction. While we did not find statistically significant heterogeneity in the impact of convictions on education attainment, further research with larger sample sizes may provide deeper insights into potential variations. Nevertheless, the overarching results suggest a potential negative influence of parental convictions on education outcomes.

Overall, our research underscores the far-reaching consequences of an individual's criminal convictions on family members. This study provides the first causal evidence of the consequences stemming from criminal convictions on family members in Brazil. By establishing a causal link, it offers compelling evidence of the far-reaching impact of such convictions, shedding light on the pervasive nature of the consequences involved. The findings highlight the need for targeted interventions and support systems to address the adverse effects of criminal convictions.

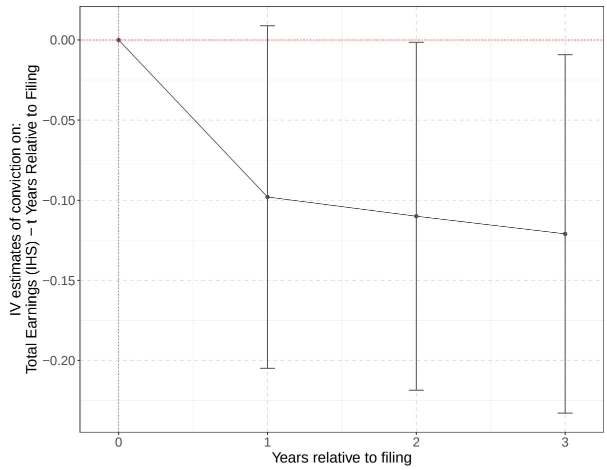
FIGURES



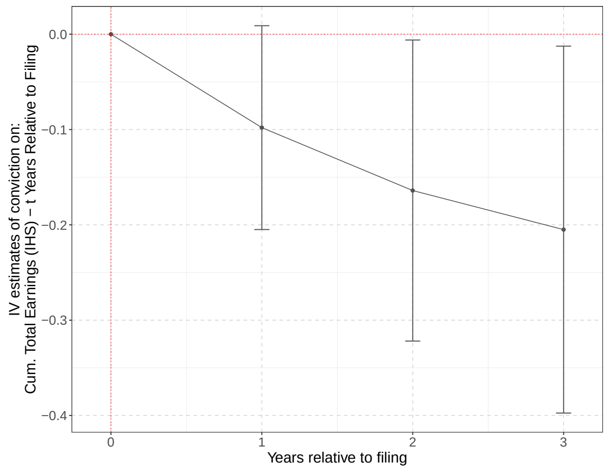
(a) Pr(Ever Worked)



(b) IHS Cum. Total Days Worked)



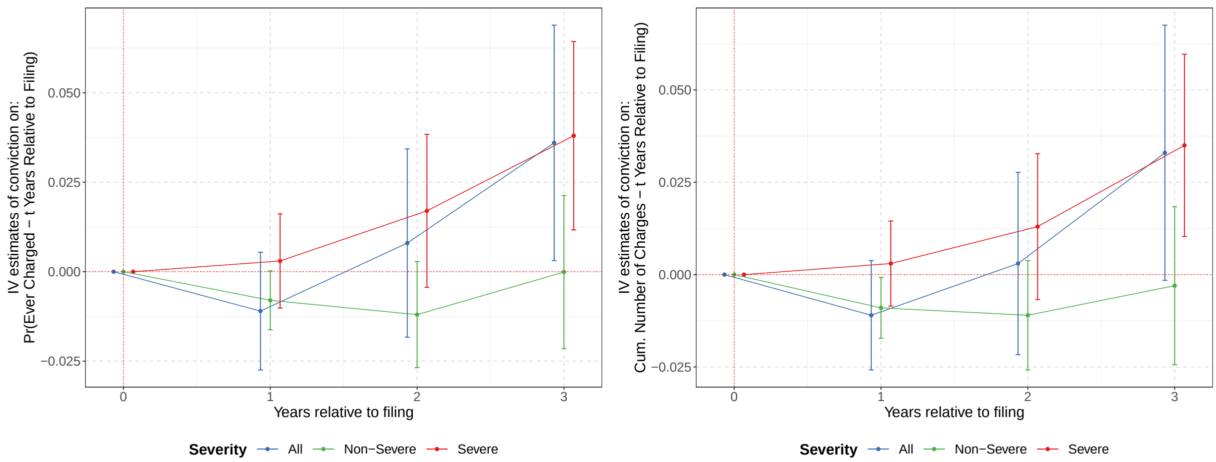
(c) IHS Total Earnings (real)



(d) IHS Cum. Total Earnings (real)

Figura 2.1 – The Effect of Conviction on Family's Labor Outcomes

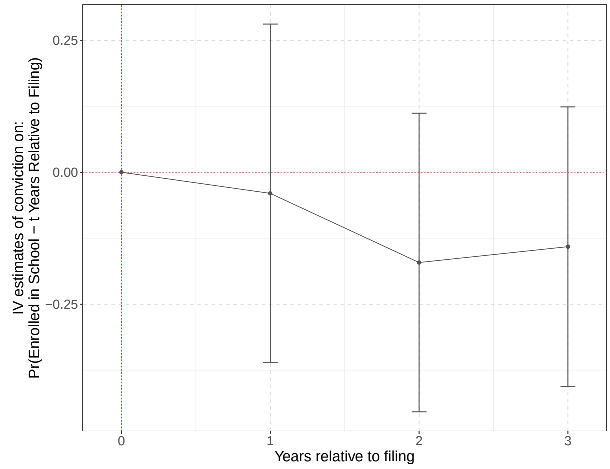
Notes: Baseline sample of criminal cases filed 2010-2017 in Brazil. Each point on the graph is the IV estimation from within-period version of our 2SLS formulation. The error bars show 90% confidence intervals.
Source: Prepared by the author.



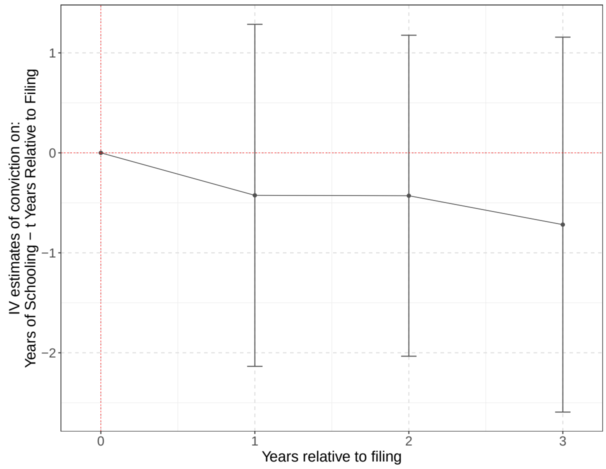
(a) Extensive Margin **(b) Intensive Margin (IHS)**

Figura 2.2 – The Effect of Conviction on Family's Future Criminal Behavior

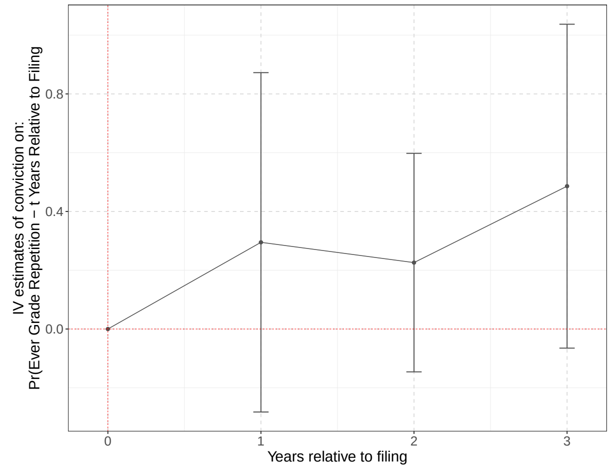
Notes: Baseline sample of criminal cases filed 2010-2017 in Brazil. Each point on the graph is the IV estimation from period-by-period version of our 2SLS formulation. The error bars show 90% confidence intervals.
Source: Prepared by the author.



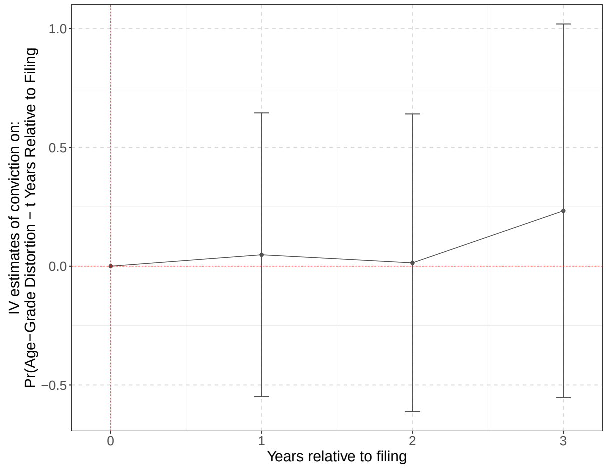
(a) Pr(Enrolled in School)



(b) Years of Schooling



(c) Pr(Ever Grade Repetition)



(d) Pr(Age-Grade Distortion)

Figura 2.3 – The Effect of Conviction on Family's Education Outcomes

Notes: Baseline sample of criminal cases filed 2010-2014 in Brazil. Each point on the graph is the IV estimation from within-period version of our 2SLS formulation. The error bars show 90% confidence intervals.
Source: Prepared by the author.

TABLES

Tabela 2.1 – Descriptive Statistics

	Overall (1)	Not convicted (2)	Convicted (3)
Relationship with Defendant			
Siblings	62.39%	47.38%	71.51%
Children	25.63%	36.82%	18.83%
Spouses	11.98%	15.80%	9.66%
Gender			
Male	54.51%	54.91%	54.26%
Female	45.49%	45.09%	45.74%
Age	21.36 (9.83)	22.37 (11.30)	20.74 (8.77)
Race			
White	54.30%	56.04%	53.24%
Black	39.86%	38.58%	40.63%
Indigenous	0.09%	0.13%	0.08%
Non Identified	5.75%	5.25%	6.05%
School			
< 9 years	6.24%	6.14%	6.31%
< 12 years	22.94%	22.70%	23.08%
>= 12 years	70.82%	71.16%	70.61%
Missing Xs	1.13 (0.91)	1.13 (0.91)	1.13 (0.91)
Observations	36,446	13,781	22,665

Notes: Baseline sample of parental criminal cases filed during 2010-2022 period. Statistics are at the individual linked to defendant level and include $36,446$ unique individuals. Column (1) reports the sample averages/proportions for the full sample. Columns (2) and (3) report the sample averages/proportions for the 'Not Convicted Defendant' and 'Convicted Defendant' sub-sample, respectively. Standard deviations are displayed in parentheses.
Source: Prepared by the author.

Tabela 2.2 – First stage estimates of conviction on courtroom stringency

	P(convicted)			
	Defendants linked to Family		All Defendants	
	(1)	(2)	(3)	(4)
Courtroom stringency	0.889*** (0.048)	0.854*** (0.069)	0.827*** (0.024)	0.821*** (0.024)
Age		-0.003*** (0.0006)		-0.003*** (0.0003)
Female		-0.033*** (0.012)		-0.068*** (0.005)
Black race		-0.007 (0.012)		0.017*** (0.005)
Indigenous race		0.062 (0.091)		0.042 (0.058)
Non identified race		-0.009 (0.024)		0.010 (0.010)
Missing Xs		-0.004 (0.012)		-0.025*** (0.008)
Worked, t-1		-0.034*** (0.011)		-0.071*** (0.008)
Worked, t-2 to t-3		0.015 (0.011)		-0.001 (0.005)
Worked, t-4 to t-5		0.004 (0.011)		-0.014*** (0.004)
<= 12 years education				-0.002 (0.011)
> 12 years education				-0.022** (0.011)
Court-Year-Subject FE	Yes	Yes	Yes	Yes
Dependant mean	0.627	0.627	0.617	0.617
Observations	17,263	13,451	56,723	56,723

Notes: Baseline sample in columns (1) - (2) consists of criminal cases filed 2010-2020 for defendants linked to family members. Columns (3) - (4) consist of criminal cases filed 2010-2020 for all defendants. The results for the years of education covariate are omitted from the Defendants linked to Family sample because of the high degree of missing information. Specifications include controls for *court x year x subject* fixed effects. Standard errors are clustered at *courtroom* level. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.
Source: Prepared by the author.

Tabela 2.3 – Testing for random assignment of cases to courtrooms

	Defendants linked to Family		All Defendants	
	Pr(Conviction) (1)	Courtroom stringency (2)	Pr(conviction) (3)	Courtroom stringency (4)
Age	-0.003***	-1.4e-5	-0.003***	5.15e-6
	(0.0006)	(0.0001)	(0.0003)	(4.46e-5)
Female	-0.038***	-0.005**	-0.072***	-0.005***
	(0.012)	(0.002)	(0.006)	(0.002)
Black race	-0.006	0.002	0.017***	0.0003
	(0.012)	(0.002)	(0.005)	(0.0009)
Indigenous race	0.082	0.024*	0.033	-0.012
	(0.093)	(0.014)	(0.059)	(0.011)
Non identified race	-0.007	0.003	0.010	-0.0003
	(0.025)	(0.005)	(0.011)	(0.002)
Missing Xs	-0.005	-0.001	-0.026***	-0.001
	(0.012)	(0.002)	(0.008)	(0.001)
Worked, t-1	-0.036***	-0.002	-0.072***	-0.0008
	(0.012)	(0.002)	(0.009)	(0.002)
Worked, t-2 to t-3	0.014	-0.001	-0.001	-0.0003
	(0.011)	(0.002)	(0.005)	(0.0009)
Worked, t-4 to t-5	0.003	-0.0008	-0.015***	-0.0006
	(0.011)	(0.002)	(0.004)	(0.0009)
<= 12 years education			-0.002	0.0002
			(0.012)	(0.002)
> 12 years education			-0.021*	0.002
			(0.011)	(0.002)
Court-Year-Subject FE	Yes	Yes	Yes	Yes
F (joint nullity), stat.	7.2020	1.4061	33.793	1.3493
F (joint nullity), p-value	1.71e-10	0.17903	1.16e-72	0.18985
Observations	13,451	13,451	56,723	56,723

Notes: Baseline sample in columns (1) - (2) consists of criminal cases filed 2010-2020 for defendants linked to family members. Columns (3) - (4) consist of criminal cases filed 2010-2020 for all defendants. The results for the years of education covariate are omitted from the Defendants linked to Family sample because of the high degree of missing information. Specifications include controls for *court x year x subject* fixed effects. Standard errors are clustered at *courtroom* level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
Source: Prepared by the author.

Tabela 2.4 – Estimates of conviction on family's labor outcomes

	1 Year	1-2 Years	1-3 Years
Panel A: Ever Worked			
OLS (all controls)	-0.027**	-0.029**	-0.030**
	(0.013)	(0.013)	(0.013)
RF (all controls)	-0.059	-0.094*	-0.070
	(0.051)	(0.050)	(0.052)
IV (no controls)	-0.078	-0.112*	-0.088
	(0.063)	(0.064)	(0.068)
IV (all controls)	-0.059	-0.093*	-0.070
	(0.050)	(0.050)	(0.052)
Dependent mean	0.406	0.469	0.508
Observations	13,191	13,191	13,191
Panel B: I.H.S Total Days Worked			
OLS (all controls)	-0.160**	-0.182**	-0.199**
	(0.077)	(0.087)	(0.089)
RF (all controls)	-0.367	-0.535	-0.510
	(0.303)	(0.325)	(0.359)
IV (no controls)	-0.483	-0.655	-0.636
	(0.382)	(0.422)	(0.473)
IV (all controls)	-0.363	-0.529	-0.503
	(0.303)	(0.325)	(0.356)
Dependent mean	2.43	3.03	3.42
Observations	13,191	13,191	13,191
Panel C: I.H.S Total Earnings (R$ 1000)			
OLS (all controls)	-0.024	-0.022	-0.019
	(0.016)	(0.015)	(0.015)
RF (all controls)	-0.099	-0.111*	-0.121*
	(0.065)	(0.067)	(0.069)
IV (no controls)	-0.134	-0.141*	-0.154*
	(0.083)	(0.083)	(0.085)
IV (all controls)	-0.098	-0.110*	-0.121*
	(0.065)	(0.066)	(0.068)
Dependent mean	0.469	0.474	0.468
Observations	12,721	12,493	12,262
Panel D: I.H.S Cum. Total Earnings (R$ 1000)			
OLS (all controls)	-0.024	-0.037	-0.040
	(0.016)	(0.023)	(0.027)
RF (all controls)	-0.099	-0.166*	-0.205*
	(0.065)	(0.097)	(0.118)
IV (no controls)	-0.134	-0.210*	-0.263*

(continued)

	1 Year	1-2 Years	1-3 Years
	(0.083)	(0.121)	(0.150)
IV (all controls)	-0.098	-0.164*	-0.205*
	(0.065)	(0.096)	(0.117)
Dependent mean	0.469	0.727	0.898
Observations	12,721	12,493	12,262

Notes: Baseline estimation sample of criminal cases filed 2010-2017. I.H.S stands for *Inverse Hyperbolic Sine*. Total Earnings and Cum. Total Earnings are expressed in units of thousands of *Reais*. Control variables used are the ones listed in Table 2.1. All specifications include *court x year x subject* fixed effects. Standard errors are clustered at *courtroom* level. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.
Source: Prepared by the author.

Tabela 2.5 – Estimates of conviction on family's criminal behavior

	1 Year	1-2 Years	1-3 Years
Panel A: Ever Criminal Charged (All Cases)			
OLS (all controls)	0.005	0.003	0.007
	(0.003)	(0.005)	(0.006)
RF (all controls)	-0.011	0.009	0.037*
	(0.010)	(0.016)	(0.020)
IV (no controls)	-0.009	0.011	0.039**
	(0.010)	(0.016)	(0.020)
IV (all controls)	-0.011	0.008	0.036*
	(0.010)	(0.016)	(0.020)
Dependent mean	0.012	0.023	0.031
Observations	13,191	13,191	13,191
Panel B: Ever Criminal Charged (Non-Severe Cases)			
OLS (all controls)	0.002	0.001	0.002
	(0.001)	(0.003)	(0.003)
RF (all controls)	-0.008*	-0.012	-0.0001
	(0.005)	(0.009)	(0.013)
IV (no controls)	-0.007	-0.010	0.001
	(0.005)	(0.009)	(0.013)
IV (all controls)	-0.008*	-0.012	-0.0001
	(0.005)	(0.009)	(0.013)
Dependent mean	0.005	0.009	0.014
Observations	13,191	13,191	13,191
Panel C: Ever Criminal Charged (Severe Cases)			
OLS (all controls)	0.002	-0.0002	0.002
	(0.003)	(0.003)	(0.004)
RF (all controls)	0.003	0.017	0.039**
	(0.008)	(0.013)	(0.015)
IV (no controls)	0.004	0.018	0.040**
	(0.008)	(0.013)	(0.016)
IV (all controls)	0.003	0.017	0.038**
	(0.008)	(0.013)	(0.016)
Dependent mean	0.007	0.013	0.017
Observations	13,191	13,191	13,191
Panel D: Cum. Num. of Charges (All Cases)			
OLS (all controls)	0.004	0.0006	0.005
	(0.003)	(0.005)	(0.006)
RF (all controls)	-0.011	0.003	0.033
	(0.009)	(0.015)	(0.021)
IV (no controls)	-0.009	0.005	0.035*

(continued)

	1 Year	1-2 Years	1-3 Years
	(0.009)	(0.015)	(0.021)
IV (all controls)	-0.011	0.003	0.033
	(0.009)	(0.015)	(0.021)
Dependent mean	0.012	0.023	0.032
Observations	13,191	13,191	13,191
Panel E: Cum. Num. of Charges (Non-Severe Cases)			
OLS (all controls)	0.002	0.002	0.002
	(0.001)	(0.003)	(0.003)
RF (all controls)	-0.009**	-0.011	-0.003
	(0.004)	(0.009)	(0.013)
IV (no controls)	-0.008*	-0.010	-0.002
	(0.004)	(0.009)	(0.013)
IV (all controls)	-0.009*	-0.011	-0.003
	(0.005)	(0.009)	(0.013)
Dependent mean	0.004	0.009	0.013
Observations	13,191	13,191	13,191
Panel F: Cum. Num. of Charges (Severe Cases)			
OLS (all controls)	0.002	-0.002	-0.0002
	(0.002)	(0.004)	(0.004)
RF (all controls)	0.003	0.013	0.035**
	(0.007)	(0.012)	(0.015)
IV (no controls)	0.004	0.014	0.036**
	(0.007)	(0.012)	(0.015)
IV (all controls)	0.003	0.013	0.035**
	(0.007)	(0.012)	(0.015)
Dependent mean	0.006	0.012	0.016
Observations	13,191	13,191	13,191

Notes: Baseline estimation sample of criminal cases filed 2010-2017. Control variables used are the ones listed in Table 2.1. All specifications include *court x year x subject* fixed effects. Standard errors are clustered at *courtroom* level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.
Source: Prepared by the author.

Tabela 2.6 – Estimates of conviction on family's education outcomes

	1 Year	1-2 Years	1-3 Years
Panel A: Pr(Enrolled in School)			
OLS (all controls)	0.041	0.026	0.008
	(0.035)	(0.031)	(0.029)
RF (all controls)	-0.034	-0.146	-0.121
	(0.166)	(0.137)	(0.133)
IV (no controls)	-0.005	-0.106	-0.067
	(0.195)	(0.186)	(0.191)
IV (all controls)	-0.040	-0.171	-0.141
	(0.195)	(0.172)	(0.161)
Dependent mean	0.868	0.838	0.801
Observations	2,182	2,182	2,182
Panel B: Years of Scholling			
OLS (all controls)	-0.152	-0.198	-0.281
	(0.148)	(0.172)	(0.191)
RF (all controls)	-0.320	-0.382	-0.588
	(0.766)	(0.863)	(0.920)
IV (no controls)	-1.43	-0.733	0.240
	(2.37)	(2.13)	(2.75)
IV (all controls)	-0.425	-0.429	-0.718
	(1.04)	(0.976)	(1.14)
Dependent mean	6.91	6.94	6.92
Observations	1,818	1,613	1,379
Panel C: Pr(Grade Repetition)			
OLS (all controls)	-0.029	-0.035	-0.010
	(0.046)	(0.035)	(0.030)
RF (all controls)	0.250	0.218	0.396**
	(0.250)	(0.188)	(0.154)
IV (no controls)	0.202	0.157	0.358
	(0.320)	(0.202)	(0.256)
IV (all controls)	0.295	0.226	0.486
	(0.351)	(0.226)	(0.335)
Dependent mean	0.128	0.122	0.115
Observations	1,698	1,504	1,285
Panel D: Pr(Age-Grade Distortion)			
OLS (all controls)	0.070	0.042	0.042
	(0.052)	(0.056)	(0.068)
RF (all controls)	0.036	0.012	0.172
	(0.273)	(0.312)	(0.346)
IV (no controls)	-0.192	-0.153	0.031

(continued)

	1 Year	1-2 Years	1-3 Years
	(0.362)	(0.370)	(0.447)
IV (all controls)	0.048	0.014	0.233
	(0.363)	(0.381)	(0.478)
Dependent mean	0.229	0.309	0.392
Observations	1,818	1,740	1,627

Notes: Baseline estimation sample of criminal cases filed 2010-2014. I.H.S stands for *Inverse Hyperbolic Sine*. Control variables used are the ones listed in Table 2.1. All specifications include *court x year x subject* fixed effects. Standard errors are clustered at *courtroom* level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Source: Prepared by the author.

Tabela 2.7 – Heterogeneity in Labor Outcomes Estimation for Family Members

		Relation to Defendant				Gender		Crime Level	
	All	Children	Spouses	Siblings	Siblings (Younger)	Male	Female	Non-Severe Crime	Severe Crime
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: P(Ever work) 1-3 Years									
IV (all controls)	-0.070	-0.006	-0.140	-0.038	0.066	-0.127	-0.057	-0.021	-0.074
	(0.052)	(0.262)	(0.198)	(0.072)	(0.120)	(0.088)	(0.095)	(0.069)	(0.101)
Dependent mean	0.508	0.498	0.450	0.523	0.558	0.504	0.514	0.526	0.473
Observations	13,191	1,325	2,249	9,617	4,307	7,491	5,700	7,695	4,019
Panel B: I.H.S Cum. Total Days Worked 1-3 Years									
IV (all controls)	-0.503	-0.150	-1.10	-0.202	0.534	-0.951	-0.244	-0.064	-0.751
	(0.356)	(1.87)	(1.34)	(0.486)	(0.820)	(0.613)	(0.651)	(0.465)	(0.690)
Dependent mean	3.42	3.31	3.03	3.53	3.73	3.37	3.50	3.56	3.14
Observations	13,191	1,325	2,249	9,617	4,307	7,491	5,700	7,695	4,019
Panel C: I.H.S Total earnings 1-3 Years									
IV (all controls)	-0.121*	0.002	-0.222	-0.050	0.005	-0.237**	0.031	-0.002	-0.213
	(0.068)	(0.250)	(0.256)	(0.082)	(0.146)	(0.119)	(0.112)	(0.081)	(0.136)
Dependent mean	0.468	0.437	0.423	0.484	0.499	0.476	0.458	0.495	0.418
Observations	12,262	1,233	2,113	8,916	3,984	6,962	5,300	7,108	3,770
Panel D: I.H.S Cum. Total Earnings 1-3 Years									
IV (all controls)	-0.205*	-0.030	-0.402	-0.080	0.086	-0.401**	0.041	-0.011	-0.373
	(0.117)	(0.468)	(0.442)	(0.149)	(0.260)	(0.203)	(0.199)	(0.143)	(0.242)
Dependent mean	0.898	0.848	0.801	0.927	0.971	0.903	0.890	0.943	0.811

Observations	12,262	1,233	2,113	8,916	3,984	6,962	5,300	7,108	3,770

Notes: Baseline estimation sample of criminal cases filed 2010:2017. I.H.S stands for *Inverse Hyperbolic Sine*. Control variables are the ones listed in Table 2.1. All specifications include *court x year x subject* fixed effects. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Source: Prepared by the author.

Tabela 2.8 – Heterogeneity in Criminal Behavior for Family Members

		Relation to Defendant				Gender		Crime Level	
	All	Children	Spouses	Siblings	Siblings (Younger)	Male	Female	Non-Severe Crime	Severe Crime
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Ever Criminal Charged (All Cases) 1-3 Years									
IV (all controls)	0.036*	0.009	0.022	0.030	0.051	0.038	0.009	0.025	0.011
	(0.020)	(0.072)	(0.067)	(0.025)	(0.050)	(0.036)	(0.013)	(0.026)	(0.038)
Dependent mean	0.031	0.038	0.027	0.032	0.034	0.051	0.006	0.030	0.037
Observations	13,191	1,325	2,249	9,617	4,307	7,491	5,700	7,695	4,019
Panel B: Ever Criminal Charged (Non-Severe Cases) 1-3 Years									
IV (all controls)	-0.0001	0.033	-0.010	-0.007	0.034	-0.017	0.007	-0.0004	-0.016
	(0.013)	(0.034)	(0.038)	(0.015)	(0.027)	(0.018)	(0.012)	(0.015)	(0.020)
Dependent mean	0.014	0.015	0.013	0.013	0.014	0.022	0.002	0.015	0.013
Observations	13,191	1,325	2,249	9,617	4,307	7,491	5,700	7,695	4,019
Panel C: Ever Criminal Charged (Severe Cases) 1-3 Years									
IV (all controls)	0.038**	-0.026	0.023	0.048**	0.045	0.047	0.0008	0.019	0.020
	(0.016)	(0.066)	(0.053)	(0.022)	(0.045)	(0.032)	(0.004)	(0.019)	(0.033)
Dependent mean	0.017	0.020	0.014	0.017	0.018	0.027	0.003	0.014	0.023
Observations	13,191	1,325	2,249	9,617	4,307	7,491	5,700	7,695	4,019
Panel D: I.H.S Cum. Num. of Charges (All cases) 1-3 Years									
IV (all controls)	0.033	0.018	-0.004	0.029	0.045	0.026	0.007	0.019	0.0009
	(0.021)	(0.066)	(0.065)	(0.027)	(0.056)	(0.036)	(0.014)	(0.025)	(0.038)
Dependent mean	0.032	0.040	0.029	0.031	0.034	0.052	0.005	0.031	0.037
Observations	13,191	1,325	2,249	9,617	4,307	7,491	5,700	7,695	4,019
Panel E: I.H.S Cum. Num. of Charges (Non-Severe Cases) 1-3 Years									

(continued)

	All	Relation to Defendant				Gender		Crime Level	
	All	Children	Spouses	Siblings	Siblings (Younger)	Male	Female	Non-Severe Crime	Severe Crime
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IV (all controls)	-0.003	0.031	-0.018	-0.008	0.028	-0.021	0.004	-0.003	-0.020
	(0.013)	(0.030)	(0.038)	(0.015)	(0.024)	(0.018)	(0.013)	(0.015)	(0.021)
Dependent mean	0.013	0.015	0.015	0.013	0.013	0.022	0.002	0.016	0.012
Observations	13,191	1,325	2,249	9,617	4,307	7,491	5,700	7,695	4,019
Panel F: I.H.S Cum. Num. of Charges (Severe Cases) 1-3 Years									
IV (all controls)	0.035**	-0.020	0.019	0.043**	0.029	0.041	0.0009	0.015	0.015
	(0.015)	(0.058)	(0.046)	(0.021)	(0.043)	(0.029)	(0.004)	(0.018)	(0.031)
Dependent mean	0.016	0.020	0.013	0.016	0.017	0.025	0.003	0.013	0.022
Observations	13,191	1,325	2,249	9,617	4,307	7,491	5,700	7,695	4,019

Notes: Baseline estimation sample of criminal cases filed 2010:2017. I.H.S stands for *Inverse Hyperbolic Sine*. Control variables are the ones listed in Table 2.1. All specifications include *court x year x subject* fixed effects. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Source: Prepared by the author.

Tabela 2.9 – Heterogeneity in Education Outcomes Estimation for Family Members

		Relation to Defendant		Gender		Crime Level	
	All	Children	Siblings	Male	Female	Non-Severe Crime	Severe Crime
	(1)	(2)	(4)	(6)	(7)	(8)	(9)
Panel A: Pr(Enrolled in School) 1-3 Years							
IV (all controls)	-0.141	0.176	-0.305	0.014	0.068	-0.135	-0.276
	(0.161)	(0.136)	(0.496)	(0.333)	(0.209)	(0.232)	(0.288)
Dependent mean	0.801	0.888	0.732	0.802	0.800	0.808	0.779
Observations	2,182	968	1,214	1,090	1,092	1,102	684
Panel B: Years of Scholling 1-3 Years							
IV (all controls)	-0.718	-0.146	-19.4	1.67	-1.21	-0.205	-1.15
	(1.14)	(0.934)	(110.5)	(3.38)	(1.06)	(1.32)	(2.42)
Dependent mean	6.92	6.27	7.67	6.77	7.08	7.18	6.99
Observations	1,379	737	642	696	683	699	418
Panel C: Pr(Grade Repetition) 1-3 Years							
IV (all controls)	0.673	0.558	6.81	0.199	0.571	0.238	0.951
	(0.553)	(0.441)	(45.3)	(0.983)	(0.590)	(0.553)	(0.750)
Dependent mean	0.344	0.301	0.388	0.386	0.303	0.335	0.341
Observations	1,426	715	711	707	719	728	443
Panel D: Pr(Age-Grade Distortion) 1-3 Years							
IV (all controls)	0.233	0.190	4.95	-0.335	0.410	-0.214	0.229
	(0.478)	(0.421)	(22.4)	(0.672)	(0.496)	(0.933)	(0.707)
Dependent mean	0.392	0.311	0.474	0.456	0.324	0.376	0.367
Observations	1,627	819	808	835	792	830	490

Notes: Baseline estimation sample of criminal cases filed 2010:2014. Control variables are the ones listed in Table 2.1. All specifications include *court x year x subject* fixed effects. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.
Source: Prepared by the author.

REFERÊNCIAS

AIZER, A.; JR, J. J. D. Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. **The Quarterly Journal of Economics**, MIT Press, v. 130, n. 2, p. 759–803, 2015.

ARTEAGA, C. Parental Incarceration and Children's Educational Attainment. **The Review of Economics and Statistics**, p. 1–45, 10 2021. ISSN 0034-6535. Disponível em: <https://doi.org/10.1162/rest_a_01129>.

BHULLER, M.; DAHL, G. B.; LØKEN, K. V.; MOGSTAD, M. **Incarceration spillovers in criminal and family networks**. [S.l.], 2018.

BHULLER, M.; DAHL, G. B.; LØKEN, K. V.; MOGSTAD, M. Intergenerational effects of incarceration. **AEA Papers and Proceedings**, v. 108, p. 234–40, May 2018. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/pandp.20181005>.

BHULLER, M.; DAHL, G. B.; LØKEN, K. V.; MOGSTAD, M. Incarceration, recidivism, and employment. **Journal of Political Economy**, v. 128, n. 4, p. 1269–1324, 2020.

BHULLER, M.; KHOURY, L.; LøKEN, K. V. **Prison, Mental Health and Family Spillovers**. [S.l.], 2021. Disponível em: <https://ideas.repec.org/p/hhs/nhheco/2021_019.html>.

CHAISEMARTIN, C. de; RAMIREZ-CUELLAR, J. At what level should one cluster standard errors in paired experiments, and in stratified experiments with small strata? Available at SSRN: https://ssrn.com/abstract=3520820 or http://dx.doi.org/10.2139/ssrn.3520820. 2023.

COHEN, A.; YANG, C. S. Judicial politics and sentencing decisions. **American Economic Journal: Economic Policy**, v. 11, n. 1, p. 160–91, 2019.

COLLINSON, R.; HUMPHRIES, J. E.; MADER, N. S.; REED, D. K.; TANNENBAUM, D. I.; DIJK, W. V. Eviction and poverty in american cities. **The Quarterly Journal of Economics**, forthcoming, 2022.

DAHL, G. B.; KOSTØL, A. R.; MOGSTAD, M. Family welfare cultures. **The Quarterly Journal of Economics**, MIT Press, v. 129, n. 4, p. 1711–1752, 2014.

DOBBIE, W.; GOLDIN, J.; YANG, C. S. The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. **American Economic Review**, v. 108, n. 2, p. 201–40, 2018.

DOBBIE, W.; GRÖNQVIST, H.; NIKNAMI, S.; PALME, M.; PRIKS, M. **The Intergenerational Effects of Parental Incarceration.** [S.l.], 2019.

DOBBIE, W.; SONG, J. Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection. **American economic review**, v. 105, n. 3, p. 1272–1311, 2015.

FARIAS, C.; SAMPAIO, B.; BRITTO, D. G. C. Conviction, employment and recidivism: Evidence from brazil. Working Paper. 2023.

FRENCH, E.; SONG, J. The effect of disability insurance receipt on labor supply. **American economic Journal: economic policy**, v. 6, n. 2, p. 291–337, 2014.

GARIN, A.; NORRIS, D. K. C. M. S.; YOTAM, M. P. E. K. R.; WEAVER, S.-T. J. The impact of incarceration on employment and earnings. Working Paper. 2022.

GREEN, D. P.; WINIK, D. Using random judge assignments to estimate the effects of incarceration and probation on recidivism among drug offenders. **Criminology**, Wiley Online Library, v. 48, n. 2, p. 357–387, 2010.

HARDING, D. J.; MORENOFF, J. D.; NGUYEN, A. P.; BUSHWAY, S. D. Imprisonment and labor market outcomes: Evidence from a natural experiment. **American Journal of Sociology**, University of Chicago Press Chicago, IL, v. 124, n. 1, p. 49–110, 2018.

HUTTUNEN, K.; KAILA, M.; KOSONEN, T.; NIX, E. Shared punishment? the impact of incarcerating fathers on child outcomes. Working paper. 2019.

HUTTUNEN, K.; KAILA, M.; NIX, E. The crime ladder: Estimating the impact of different punishments on defendant outcomes. Working paper. 2020.

JR, J. J. D. Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care. **Journal of political Economy**, The University of Chicago Press, v. 116, n. 4, p. 746–770, 2008.

KLING, J. R. Incarceration length, employment, and earnings. **American Economic Review**, v. 96, n. 3, p. 863–876, 2006.

LOEFFLER, C. E. Does imprisonment alter the life course? evidence on crime and employment from a natural experiment. **Criminology**, Wiley Online Library, v. 51, n. 1, p. 137–166, 2013.

MAESTAS, N.; MULLEN, K. J.; STRAND, A. Does disability insurance receipt discourage work? using examiner assignment to estimate causal effects of ssdi receipt. **American economic review**, v. 103, n. 5, p. 1797–1829, 2013.

MUELLER-SMITH, M. The criminal and labor market impacts of incarceration. **Unpublished Working Paper**, v. 18, 2015.

NORRIS, S.; PECENCO, M.; WEAVER, J. The effects of parental and sibling incarceration: Evidence from ohio. **American Economic Review**, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, v. 111, n. 9, p. 2926–2963, 2021.

OLIVEIRA, L. F. B. d.; SOARES, S. S. D. **O impacto do Programa Bolsa Família sobre a repetência: resultados a partir do cadastro único, projeto frequência e censo escolar**. [S.l.], 2013. Disponível em: <https://ideas.repec.org/p/ipe/ipetds/1814.html>.

TELLA, R. D.; SCHARGRODSKY, E. Criminal recidivism after prison and electronic monitoring. **Journal of political Economy**, University of Chicago Press Chicago, IL, v. 121, n. 1, p. 28–73, 2013.