Formal Employment and Organized Crime: Regression Discontinuity Evidence from Colombia

Gaurav Khanna* Carlos Medina†
Anant Nyshadham‡ Jorge Tamayo§

January 4, 2021

Abstract

Canonical models of entry into crime emphasize sorting on economic incentives. Yet, causal evidence of sorting into criminal occupations in response to individual-level incentives is limited. We link administrative socioeconomic microdata with the universe of arrests in Medellín over a decade. We exploit exogenous variation in formal-sector employment around a socioeconomic-score cutoff, below which individuals receive benefits if not formally employed, to test whether disincentivizing formal-sector employment induces crime. Regression discontinuity estimates show this policy generated reductions in formal-sector employment and a corresponding spike in arrests associated with criminal enterprise activity. Consistent with an occupational choice interpretation, we find no effects on crimes unlikely to be associated with organized entities, such as crimes of impulse or opportunity. Effects on arrests are strongest in neighborhoods with more opportunities to join criminal enterprises.

Keywords: *occupational choice, criminal enterprise, crime, informality, Colombia*
JEL Codes: J24, J46, K42

*University of California – San Diego, gakhanna@ucsd.edu, gauravkhanna.info
†Banco de la Republica de Colombia, cmedindu@banrep.gov.co
‡University of Michigan & NBER, nyshadha@umich.edu, anantnyshadham.com
§Harvard University, Harvard Business School, jtamayo@hbs.edu, jorge-tamayo.com
We thank seminar participants at UCSD, Michigan (H2D2), Rosario (Bogota), SoCCAM, PacDev, Empirical Studies of Conflict, Barcelona GSE Summer Forum, Economics of Risky Behavior (Bologna), Colombian Central Bank (Bogota), Minnesota (MWIEDC), Hawaii, American Economic Association (ASSAs), and the Inter American Development Bank (AL CAPONE) for feedback, and Achyuta Adhvaryu, Jorge Aguero, Prashant Bharadwaj, Chris Blattman, Michael Clemens, Gordon Dahl, Rafael DiTella, Gordon Hanson, Mauricio Romero, Emilia Tjernstrom, Juan Vargas, Mauricio Villamizar and Jeff Weaver for insightful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research, and do not commit Banco de la República or its Board of Directors.

1 Introduction

Many countries, particularly across the developing world and in much of Latin America, are plagued by coincident high degrees of informality in the labor market and criminal activity often controlled by organized enterprises (Arteaga, 2019; Brown and Velasquez, 2017; Buonanno and Vargas, 2018; Chimeli and Soares, 2017; Dell et al., 2018; DiTella et al., 2010; Sviatschi, 2018). Classic models of entry into crime contend that individuals rationally weigh the expected costs and benefits of engaging in criminal activity relative to alternatives: primarily legitimate employment in the labor market (Becker, 1968; Ehrlich, 1973). However, the empirical evidence is limited on whether formal employment and criminal enterprise activity are causally linked by way of occupational sorting decisions at the individual level.

We use rich administrative data between 2002 and 2013 from Medellín, Colombia to test the relationship between formal employment and participation in crime at the individual level. In this empirical context, where informality is common and criminal enterprise activity abounds, financially dissuading individuals from engaging in formal employment could drive some to crime as their most lucrative option outside of the formal sector. Exploiting a discontinuity in the incentives to remain informal in Colombia, we investigate individual-level occupational sorting into crime.

The Colombian government provides health benefits to all residents of households that have a socio-economic score (known as the *Sisben* score) below a certain threshold. Formal employment of any member reduces the likelihood that the household qualifies for this program, raising the relative benefits to informality. Those formally employed are automatically taxed a fraction of their wages to avail of comparable benefits. Accordingly, the usual benefits to formal employment (e.g., higher wages, job security, legal protections) are at least partially offset by the increased cost of health care coverage for those below the cutoff, who would be eligible for free coverage by the government if they were not formally employed. The importance of these incentives in our context is emphasized by the near complete health care coverage in the population, despite costs representing large proportions of income for many households.

Using a regression discontinuity design, we find that the policy induced a roughly 4 percentage point lower formal employment rate at the margin, consistent with estimates from previous studies.[1] These same individuals are arrested at higher rates for crimes likely associated with criminal enterprises (LACE). At the RD cutoff we find a 0.45 percentage point rise in LACE violent crimes (e.g., firearms trafficking, homicide), a roughly 0.66 percentage point rise in LACE property crimes (e.g., car theft), and a less precisely estimated 0.1 percentage point rise in LACE drug crimes (e.g., cocaine and heroin distribution).[2] Importantly, offenses not likely associated with economic motivations and organized criminal enterprises (e.g., rape and marijuana consumption) do not show significant increases at the cutoff, allowing us to rule out many alternative theories.[3]

At the cutoff, the program encouraged informality. High-crime environments like Medellín have an informal market that contains many lucrative opportunities to be "employed" by criminal enterprises. Additional results show that impacts on LACE crimes are strongest in neighborhoods known to have the highest gang activity at baseline. Our results suggest that increases in formal employment can lead to reductions in criminal enterprise activities. Our magnitudes are in line with related studies, as we measure an economically meaningful 3.1% increase in arrests for every 1 percentage point fall in formal employment.[4]

Our contributions lie in validating economic models of occupational choice and entry into crime (Becker, 1968; Ehrlich, 1973).[5] We leverage individual-level variation in employment incentives and rich administrative data to establish a causal relationship between formal employment and participation in criminal enterprise activity. Such evidence has proven difficult to find in a literature that has mostly relied on aggregate shocks, such as local recessions or trade-shocks. In contrast, the variation we leverage compares otherwise similar individuals with no change to their aggregate environment.

Recent studies have highlighted how unemployment shocks, job loss and employment restrictions lead to increases in criminal activity (Bell et al., 2018; Bennett and Ouazad, 2018;

[1] When evaluating the effect on the entire country using a different research design, Camacho et al. (2014) find that the program led to a 4 percentage point decrease in formal employment.

[2] Additional results in which joint outcomes of non-formal employment *and* criminal arrests are studied confirm that those leaving formal work and those arrested are the same. The results for LACE drug crimes are significant when we simultaneously measure both non-formal employment and arrests as an outcome.

[3] For instance, if social benefits induce risky behavior it should increase non-LACE crimes as well.

[4] These estimates are in the range of elasticities we derive from mass-layoffs in Medellín (Khanna et al., 2020). In Section 8 we compare our elasticities to prior work.

[5] More recent studies have explored similar mechanisms in related phenomenon, like participation in civil conflict (Dal Bó and Dal Bó, 2011; Dube and Vargas, 2013).

Khanna et al., 2020; Pinotti, 2017; Rose, 2019; Schnepel, 2018). We build upon this small set of recent studies by documenting occupational sorting as a result of exogenous variation in exposure to a tax on formal wages. Job losses and structurally imposed employment restrictions may induce effects on depression, subsequent job search, and social stigma that are less likely in our setting. We stress the importance of distinguishing between different types of crime, as some are more likely reflective of occupational sorting into criminal enterprise activity (e.g., firearms trafficking and heroin distribution) whereas others are more likely to be crimes of impulse or opportunity (e.g., rape and drug consumption). In doing so, we establish a falsification test to rule out alternative mechanisms that have little to do with occupational choice.

Additionally, there are fewer studies in the developing world, as many look at the US, the UK or Scandinavian countries from which data are more readily available (Bhuller et al., 2018; Hjalmarsson and Lindquist, 2019). We study a high-crime environment similar to most parts of the developing world and, in particular, a city with a significant presence of organized crime, which has been shown to have particularly detrimental effects on growth and development (Alesina et al., 2017; Blattman et al., 2018; Melnikov et al., 2019; Velasquez, 2020), and broader consequences for child development (Arteaga, 2019). We build upon recent evidence from important high-crime environments in Latin America that leverages area-based variation from trade-shocks (Dell et al., 2018; Dix-Carneiro et al., 2018), or district-level unemployment (Buonanno and Vargas, 2018; Cortes et al., 2016), by documenting occupational sorting in response to incentives at the individual level.[6] Finally, we highlight an unintended, adverse consequence of a welfare policy, contributing to previous work on the interaction between public sector interventions and crime (Agan and Starr, 2018; Chimeli and Soares, 2017; Chioda et al., 2016; Sviatschi, 2020; Yang, 2008).[7]

[6]This literature also shows that local trade shocks also affect public goods provisioning, inequality and policing, suggesting that the general equilibrium effects may be substantial (Dix-Carneiro et al., 2018; Feler and Senses, 2017). Indeed, Dix-Carneiro et al. (2018) extensively discuss the various channels through which such aggregate shocks may affect crime.

[7]Related work studies how elected officials may engage in criminal activity (Ferraz and Finan, 2008, 2011; Olken and Pande, 2012), and how multiple prices for public programs lead to distortions (Barnwal, 2018).

2 Background

2.1 Previous Studies

Even though models of criminal activity are based on individual behavior, previous studies have often tested these models using aggregate area-based relationships like unemployment shocks (Agan and Makowsky, 2018; Cornwell and Trumbull, 1994; Entorf, 2000; Gould et al., 2002; Machin and Meghir, 2004). Area-based relationships are meaningful and policy relevant as they inform how to broadly target crime deterrence strategies.[8] Yet, variation at the individual level is likely to produce different estimates than those that rely on aggregate shocks.

For instance, unemployment at the regional level reduces the returns to criminal activity (i.e., lowers the resources available to expropriate and is correlated with fewer potential victims in the area (Mustard, 2010)). General equilibrium effects in which a new stock of criminals may crowd-out others, and neighborhood and peer effects both within and across neighborhoods might affect the relationship between area-based employment and choices to engage in crime (Cullen et al., 2006; Dustmann and Damm, 2014; Ihlanfeldt, 2007; Kling et al., 2005, 2007).[9] Additionally, detection rates of crime outcomes may defer as local resources change. Lastly, economic shocks may have wide-ranging spillovers (Gathmann et al., 2020), as economic activity and high-income individuals may leave areas with high or increasing crime (Cullen et al., 2005; Cullen and Levitt, 1999; Greenbaum and Tita, 2004), further affecting the association between crime and employment observed at the aggregate level. In our work, we are interested in the individual-level decision to engage in criminal activity as it informs some root causes of why some choose a life of crime.

Some of the best evidence of individual-level decisions related to crime comes from experiments that raise the human capital of individuals (Blattman and Annan, 2015; Heller, 2014; Schochet et al., 2008). We complement this evidence on how changes to human capital affect the returns to both standard employment and criminal activity, by examining the occupational choice between legitimate and illegitimate activity as relative incentives are changed. Many

[8] As shown by Foley (2011); Karin (2005); Lin (2008); Raphael and Winter-Ember (2001).

[9] Fella and Gallipoli (2014) find that general equilibrium effects explain a substantial portion of the relationship between crime and schooling.

studies that attempt to examine individual-level occupational choices between legitimate and criminal activities rely on associations conditional on rich sets of observables (Freeman, 1999), as plausibly exogenous variation is challenging to find.

However, it may be difficult to account for unobservables that would determine both employment and crime. For instance, factors like high discount rates determine both crime and job-search (DellaVigna and Paserman, 2005), whereas childhood shocks and decisions may affect both adult employment and crime (Doyle, 2008, 2007; Lochner and Moretti, 2004; Sviatschi, 2018). Reverse causality leads to upward bias as employers are less likely to prefer individuals that may display attributes correlated with criminal behavior (Grogger, 1995; Kling, 2006; Lott, 1992). Unemployment rates can affect the number of victims even if there are no new criminals: employed persons may have resources that are targeted, or the unemployed may be in the crossfire or suffer substance abuse. Last, many studies depend on self-reported crime that may under-measure the occurrence of criminal activity, or homicides and victimization rates which capture the likelihood of being a victim rather than a criminal (Freeman, 1999).

We overcome each of these issues raised by previous researchers in examining the relationship between formal employment and organized crime in Medellín, Colombia. First, we link two sources of administrative data at the individual level: the universe of arrests and the pre-arrest socio-economic characteristics of citizens, overcoming measurement issues in self-reported criminal activity and aggregate area-based measures of crime. Administrative data allow us to leverage individual-level variation and focus on demographics more likely to be affected. Next, we exploit quasi-experimental variation in the relative gains to informality derived from a social benefits program that requires individuals to be outside the formal sector to be eligible. Rather than associations conditional on observables, we use exogenous variation in financial incentives to isolate the individual-level relationship between employment and crime. Last, our data allow us to distinguish between different types of crime and conduct falsification tests by comparing the impacts on crimes likely associated with criminal enterprises (LACE) to the impacts on other, more idiosyncratic crimes of impulse and opportunity.

2.2 Crime in Medellín

Located in the north-western region of Colombia, Medellín is the second largest city after the capital, Bogota. The urban zone consists of 249 neighborhoods, divided into 21 (*comunas*), 5 of which are semi-rural townships (*corregimientos*). Economic incentives are intricately tied to criminal participation in Medellín (Khanna et al., 2020).

Although Colombian violence has traditionally been high, the emergence of drug cartels in the late 1970s and early 1980s, fueled the emergence of organized crime to support illegal businesses, and guerrilla or paramilitary groups to care for the entire production chain. From the mid 1980s to early 1990s, homicide rates rose rapidly driven by the boom of cartels, paramilitaries, and local gangs. Over the period covered by our data, Medellín was one of the most violent cities in the world (CCSPJP, 2009), placing our analysis among a handful that study motivations behind joining organized crime in high-crime environments. The high homicide rates are a result of fights among urban militias, local gangs, drug cartels, criminal bands, and paramilitaries based in surrounding areas.[10] Demobilized militias continue to be involved in extortion and trafficking, given their experience with guns and avoiding police (Rozema, 2018).

There are two features of the homicide rate that are pertinent for our analysis. First, it is predominantly male. In 2002, the first year of our data, the male homicide rate was 184 per 100,000 whereas the female homicide rate was about 12, less than one-tenth the rate of males. Over the entire sample period (2005-13), 12% of all males (across all age groups) were at some point arrested, while the arrest rate for females was only 1%. Second, youth are far more likely to be involved than other age groups. Approximately, 63% of all first arrests are between 13 and 26. Younger individuals are more likely to be engaged in drug trafficking and consumption, whereas slightly older individuals are involved in violent crimes (homicides, extortions, and kidnapping), and the oldest still are involved in property crime. Irrespective of type of crime, however, arrest rates peak within the 13 to 26 age window depicted in Figure A2.

In ongoing research, Blattman et al. (2018) document Medellín's criminal world as hundreds of well-defined street gangs (*combos*) controlling local territory, organized into hierarchical relationships of supply, and protected by the *razones* at the top of the hierarchy. They confirm

[10] *Operacion Orion*, followed by the demobilization of paramilitary forces led to a sharp decline in homicides, as the military clamped down on urban militias (Medina and Tamayo, 2011).

that gangs are mainly profit-seeking organizations, earning money from protection, coercive services such as debt collection and drug sales. Anthropological studies and in person interviews show that economic incentives (such as the focus of our study) drive young men in Medellín to join organized crime (Baird, 2011). As many respondents highlight, the reason to join crime is mostly "economic" or for a profitable career.[11] Knowing this, groups actively recruit idle youth that are *amurrao* (local slang: 'sitting on the wall') and without a formal sector job.

An interview with El Mono (*p191*) documents the recruitment process: "*those guys would hang out around here and be nice to me and say 'come over here, have a bit of money'.*" Having a formal sector job means that one is not "hanging around the neighborhood" when the gangs come to recruit. A desirable outside option would be a job with benefits and social security, yet those with formal sector jobs pay extortion fees to gangs.[12] Indeed, the options are often presented as an occupational choice: "*are you gonna work [for the gang] or do a normal job?*"[13]

Often remunerations for gang members are higher than legitamate jobs for those with similar levels of education (Doyle, 2016). New recruits are employed to run guns (*carritos*), before transitioning to extortion and trafficking. Blattman et al. (2018) estimate that foot soldiers of the combos receive well above national minimum wage whereas *combo* leaders earnings "put them in the top 10% of income earners in the city." These anecdotes are consistent with our hypothesis: better benefits for informal workers discourage young men from joining the formal sector, which in turn leads many to be recruited by criminal enterprises.[14]

For our sample of young men in the bandwidth of analysis, 21.5% were arrested over the period of study – 11.1% for drug crimes, 5.6% for property crime, and 4.8% for violent crimes. These numbers are high relative to most contexts, but are representative of cities in Latin America. The US has an incarceration rate more than six times the typical OECD nation, where one in ten youths from a low-income family may join a gang, 60% of crimes are committed by offenders under the age of 30, and 72% by males (Kearney et al., 2014). Accordingly, in some regards, arrests in our context are similar to high-crime regions in many parts of the

[11]See *interview with Gato, p264* and *interview with Armando, p197.*

[12]See *interview with El Peludo, p184.*

[13] See *interview with Notes, p193*

[14]Internalizing this trade-off, during the demobilization of militias in the mid-2000s, many were encouraged to join the formal sector, given identity cards and medical cards (Rozema, 2018). Yet, this disparity in costs across social benefit regimes, discouraged formal sector re-integration.

developing world, and especially Latin America (Brown and Velasquez, 2017; Sviatschi, 2018).

2.3 Access to Health Benefits

In 1993, *Law 100* established two tiers of health insurance: the Contributive Regime (CR) and the Subsidized Regime (SR). The CR covers formal workers with a comprehensive set of health services that includes nearly all of the most common illnesses. The SR covers the families of the poorest informal workers and unemployed with a plan that initially covered fewer illnesses than CR, but was expanded to cover the same benefits.[15] Formal workers and employers fund workers' insurance premiums for coverage by the CR. Between the 1993 reform and 1998, insurance coverage under both grew from 20% to 60%. In 2005, SR was expanded and takeup reached 1.1 million people in Medellín, alone. By 2013, 96% of Colombians were covered, with more than half qualifying under SR (Lamprea and Garcia, 2016).

Colombian employers are required by law to enroll all their employees in a Health Promoting Company, which gives them access to health insurance under the CR. Formal employees pay up to 12.5% of their wage, and cross-subsidize informal workers in SR.[16] Authorities initially expected the formal sector population to rise and cover costs for SR. But the SR grew faster than the CR population, in part due to the generous nature of the SR (Lamprea and Garcia, 2016). Costs to formal workers rose substantially after the 1993 reforms, with strong evidence that such costs discouraged formal sector employment (Kugler and Kugler, 2009).

Self-employed workers are allowed to enroll in the CR themselves by paying a monthly fixed amount based on a percentage of the monthly minimum wage. Unemployed or inactive individuals (and informal workers) can either get health insurance as the self-employed do through the CR, or apply for access to the SR. Individuals not covered by the CR or the SR use public hospitals, and are charged fees for both medicines and services.[17]

[15] In 2008, the Constitutional Court ordered that the basket of health services covered under SR be equal to that of CR. CR was slightly more comprehensive in the early years, which could have theoretically weakened our first-stage on being formal. Nevertheless, we find strong impacts on both SR enrollment and formal employment as presented below. At the beginning of our sample period, the subsidized program covered nearly 60% of health services that the full program covered – this fraction increased steadily to 100% of services.

[16] Formal sector workers make up about 54% of the urban labor force and pay 4% of their monthly wage for enrollment in the CR, while the employer pays the other 8.5%. This implies that effectively employees may bear a burden somewhere between 4 and 12.5% of their monthly wage depending on their bargaining power. Formal workers pay 1.5% of their salary to cover informal workers in SR.

[17] For instance, to the best of our knowledge, there are no separate lines at hospitals, or differences in wait

To target the SR, roughly 70 percent of the poorest households in the country were interviewed between 1994 and 2003, and a welfare index (*Sisben* score) was calculated using a confidential formula based on respondent characteristics, incomes and assets, disability, education, and housing. Only households with a *Sisben* score below a certain cutoff and not formally employed were eligible to become beneficiaries of the SR.[18] Other public programs use the *Sisben* score, but the SR *Sisben* cutoff did not coincide with other major interventions, at the eligibility cutoff of *Sisben* during the study period.[19] The SR health program is by far the largest that has eligibility determined by the *Sisben* score.[20]

2.4 Incentives for Informality

Between 2005 and 2013, informal workers made up 46% of the urban labor force in Medellin.[21] Effectively, financial incentives embodied in the health coverage options switch from potentially promoting formal employment above the cutoff due to a partial defrayal of the costs of healthcare by the employer, to strongly discouraging formal employment below the cutoff due to a significantly more enticing full defrayal of these large costs by the government for individuals who are not formally employed. Near complete health care coverage in the population, despite costs representing large proportions of income, reveals the importance of these incentives.

That this policy led to a fall in formal-sector employment has been documented in both the academic literature and public discourse. The Minister of Social Protection, in a news article in *Presidencia de la Republica* (February, 2006), claimed that the people's valuation of SR was so high that it discouraged formal employment. Studying the effects on the entire country, Camacho et al. (2014) use individual-level data and control for both region and time fixed effects to show that informality increased by 4 percentage points as SR was rolled out

times for appointments under the two regimes.

[18] Households keep their Sisben score until it is updated by the government (expected to take place about every five years). In this case, the government updates the Sisben survey and score for the entire country.

[19] See www.sisben.gov.co/Paginas/Noticias/Puntos-de-corte.aspx for programs by *Sisben 3* cutoff. While the *Sisben* cutoff for SR enrollment may differ across counties, there is only one cutoff for the entirety of Medellín, and it did not coincide with other programs.

[20] The share of the SR in the total budget accounts to nearly 2% of the GDP, while all other programs sought to reduce poverty represent less than 0.4% of GDP.

[21] Among informal workers, around 60% were own account workers, 20.5% were private sector employees, 7.8% were domestic workers, 7.7% self-employed workers, and the rest were laborers, family workers, and unpaid workers. There is a stark education gradient, where more education is associated with a higher likelihood of formal employment (Medina et al., 2013).

across the country. This is a combination of workers dropping out of the formal sector, but also fewer youth joining the formal sector over time (Lamprea and Garcia, 2016). Recognizing these adverse effects on formal employment, the government drastically lowered the costs of being enrolled in CR right at the end of our study period, when Law 1607 was enacted. Not surprisingly, this led to a significant increase in formal sector employment (Bernal et al., 2017; Fernández and Villar, 2017; Kugler et al., 2017; Morales and Medina, 2017).

Since the *Sisben* score is calculated at the family rather than individual level, older family members have reason to discourage youth within the family from joining the formal labor force for fear of losing access to benefits.[22] The Ministry of Health maintains a census of all those enrolled in both CR and SR, and if someone in the household becomes formal, and so enrolls in CR, they and all dependents (including adult dependents) are made ineligible for SR.[23]

Large families stay informal in the hope of retaining benefits (Joumard and Londono, 2013). Similarly interviews in Baird (2011) highlight how being involved in crime can sometimes be a 'family decision.' This is confirmed by our other work in the same context that show spillovers in criminal activity across generations within the family (Khanna et al., 2020). Indeed, Santamaria et al. (2008) find that half of all SR recipients indicated that they would not switch to formal employment as it would mean losing benefits. These effects are not restricted to men, as women's formal-sector participation also decreased in response to SR (Gaviria et al., 2007). Yet, we find that dis-employment effects on men are about four times larger than on women, consistent with the hypothesis that men have a lucrative alternative outside the formal sector: criminal enterprises.

We leverage the fact that the costs of accessing these benefits change discontinuously at the *Sisben* cutoff. Indeed, as most individuals are covered by one healthcare regime or the other,[24] almost everyone has access to benefits on either side of the cutoff by the end of this period.[25] Yet, on one side of the cutoff these benefits are free only if you are not formally employed. The primary driving variation, therefore, is that being outside the formal sector allows you to not

[22] By Article 21, Decree 2353 of 205, the *Sisben* score is determined at the family level.

[23] This administrative database is updated continuously. Additionally, becoming formal will also increase your *Sisben* score.

[24] By 2013 the coverage is 96% (Lamprea and Garcia, 2016).

[25] If anything, in the early years, CR was slightly more comprehensive, likely hindering us from finding a strong first-stage on employment.

pay for benefits on one side of the cutoff.

3 Data

Administrative data allow us to identify the relationship between incentives for informality and participation in criminal enterprise. We do not need to rely on self-reported or aggregate victim counts. As our data is at the individual-level, we isolate vulnerable demographics (young men), and test both employment outcomes and crime. Additionally, detailed information on the types of crime allow us to isolate mechanisms.

We combine two sources of data at the individual level using national identification numbers and dates of birth. One source is from successive *Sisben* surveys of the Medellín population in 2002 (baseline *Sisben I*), 2005 (*Sisben II*) and 2009-2010 (*Sisben III*). The *Sisben* dataset consists of cross sections from censuses of the poor, and we match household records across the three waves.[26] The second source is the census of individuals arrested between 2002-13 from the Judicial Police Sectional of the National Police Department.

We follow the literature and restrict our analysis to data on first arrests. Repeat arrests are excluded as time spent under incarceration and the length of sentencing may be endogenous to other characteristics.[27] Indeed, first arrests most closely map to the first decision node between legal and illegal activities. Once captured a criminal career begins, with subsequent decisions to repeat, escalate, or exit the criminal sector based on many factors we do not observe (including prison sentences). Accordingly, subsequent criminal behavior is outside the scope of this study.

For similar reasons, we follow recent studies (Gronqvist, 2017; Kling et al., 2005) in focusing on young men in our analysis. Our primary sample is between 21 and 26 years old in the last year of our arrest data, or between 13 and 26 for the entire period (2005-2013) of study, capturing more than 63% of all first arrests (as shown in Figure A2). Only about 4% of arrests are below the age of 16, and underage arrests are less likely to lead to prison sentences, and more likely to be short stays in special youth shelters (Guarin et al., 2013; Ibanez et al., 2017). Expanding the age criteria produces similar results.

[26]These are municipality censuses of people living in the three poorest socioeconomic strata. In large cities, like Medellín, it amounts to 65-80% of the population.

[27]We show that our results are robust to including repeat arrests.

Table 1: Summary Statistics in 2002

Variable	*Complete Sample* Mean	Std. Dev.	*Males* Mean	Std. Dev.
Individual Characteristics				
Male	0.490	0.500	1.000	0.000
Subsidized Regime	0.319	0.466	0.312	0.463
Contributive Regime	0.228	0.420	0.222	0.416
Age 10-15	0.105	0.306	0.109	0.311
Age 15-20	0.105	0.306	0.110	0.313
Age 20-25	0.089	0.285	0.093	0.290
Age 25-30	0.068	0.251	0.068	0.251
Ever Arrested	0.062	0.242	0.114	0.318
Household Head (HH) Characteristics				
Female	0.387	0.487	0.308	0.462
Employed	0.628	0.483	0.643	0.479
Unemployed	0.106	0.308	0.107	0.309
Married	0.345	0.475	0.377	0.485
Attending School	0.009	0.097	0.008	0.089
Has CR	0.219	0.413	0.214	0.410
Age	43.237	14.302	43.869	14.159
Years of Education	4.542	2.451	4.480	2.454
Owns House	0.314	0.464	0.327	0.469
Sisben Stratum 1	0.271	0.444	0.273	0.446
Sisben Stratrum 2	0.620	0.485	0.620	0.485
Sisben score	45.707	9.901	45.716	9.908
Number of members in household	4.090	1.709	4.215	1.709
N	1,161,446		568,923	

Summary tabulations using *Sisben* I survey, conducted in the year 2002, and police arrests data.

Of the individuals arrested more than once during the observation period for any crime (not just LACE), 40% are first arrested before the age of 27. At the same time, while incarcerated, individuals would not be able to be arrested for additional crimes and would, therefore, have lower measured propensities to be engaged in new criminal activity. Older individuals may have been arrested in their youth (or currently still be incarcerated) but as our crime data only begins in the early 2000s, we do not have their entire criminal history, and would miss their youth arrest. As such, we exclude older men. Focusing on ages when arrest rates peak reduces these concerns regarding the measurement of criminality, and allows us to emphasize the period when young men first make choices between crime and other jobs in Medellín (Doyle, 2016).

Figure A1 describes the timeline of our data. We use the 2002 *Sisben* as our baseline to create our running variable and predict eligibility for SR.[28] We test for SR enrollment in the 2005 *Sisben*, and for employment status and incomes in the 2009 *Sisben*. We then follow the criminal histories of young men aged 21 to 26 in 2013, between 2005 (after we have a measure of SR enrollment from the second *Sisben*) and 2013.

Table 1 presents the 2002 baseline summary statistics of the complete *Sisben* survey and for the subsample of male youth only.[29] The arrests data include a detailed description of the person arrested (national identification number and date of birth), type of crime (e.g., homicide, rape, motor vehicle theft, etc.), the precise penal code article associated with the crime, the date of arrest, the location of arrest, and a police flag for whether the officer knew the perpetrator to be gang affiliated.

3.1 Classifying Crimes

We classify the crimes into three categories – violent, property, and drug crimes – based on the US Bureau of Justice Statistics' classifications in the Sourcebook of Criminal Justice Statistics (BJS, 1994).[30] We also worked closely with senior police officials in Medellín to divide our crimes into those "likely associated with criminal enterprises" (LACE) and those more likely to reflect impulse, passion, or opportunity (non-LACE). For about 30% of our data, the police used a system that flagged the arrest with whether the individual was known to be part of an organized criminal enterprise or not, as well as information on the specific organization to which the individual belonged. This organizational affiliation was based on extensive police intelligence and follow up interviews. The data span 284 street gangs, urban militias, *narcotraficante* (drug distributors linked to gangs), and other organized criminal entities.

[28] The formula to compute the Sisben score and the eligibility cutoff varies across the waves (I, II, and III). One might be additionally interested in the 2009 SR enrollment as it also falls within the window of crime outcomes, however the 2009 enrollment depends on the 2005 Sisben score and cutoff, which was collected three years after we begin measuring our crime outcomes. Since our crime data begins in 2002, we use the 2002 Sisben score as our running variable and so measure SR enrollment in 2005.

[29] The SR status is established based on the previously computed Sisben score, from the semi-decadal Sisben municipality census of the 70% of the poorest population. After a Sisben survey, it takes around one or two years to get the new Sisben score and household eligibility. Accordingly, we use the lagged Sisben score as the running variable for our analysis.

[30] If an individual was first arrested for violent crime and later for property crime, they show up as an arrest for violent crime.

The police discontinued this system after a period of time. As the gang-flag system was not available for the entirety of our period, we use the patterns revealed over the subset of arrests for which the flag is available to classify the full sample of arrests. Police officials advised us that the best way to classify arrests is along two dimensions: (1) the crime and (2) the location. Accordingly, in our main analysis, we classify a crime as likely associated with criminal enterprises (LACE) if more than 30% of recorded arrests for that crime had the gang flag (our results are similar when using the median as the cutoff). As a result, for example, we classify homicide arrests as violent LACE, and rape or domestic violence as violent non-LACE.[31]

In robustness checks, we use a method that relies on the association between these crimes and historically high-gang neighborhoods. In this alternative definition, we classify crimes as LACE if they are more likely (above the median) to list any of these high-gang neighborhoods as the location of arrest.[32] While neither classification is perfect, the robustness across classification methods helps to validate the exercise.[33]

In Appendix Table A1, we categorize the 25 (of 103) most prevalent crimes under each classification method. These data-driven methods line up with our priors on types of crime: homicides, motor vehicle theft, extortion, kidnapping, break-ins, and the manufacturing, delivery and trafficking of drugs fall under LACE crimes. The remaining crimes are often thought of as crimes of impulse or opportunity (like rape, simple assault, and drug consumption). Indeed, we can distinguish between nuanced details – such as trafficking cocaine (LACE) vs. consuming marijuana (non-LACE). The advantage of these classification approaches is that they are purely data driven. Additionally, they may speak to the types of activities that gangs in Medellín engage in: for example, they are more likely to engage in car theft than identity theft.

[31]Gang rape gets classified as a gang crime. Our police contacts also describe how burglars are imbibed into gangs based on their work territories and would find it difficult to be a burglar without being a part of the gang.

[32]Our results are similar when using other cutoffs from 30% to 60% of arrests.

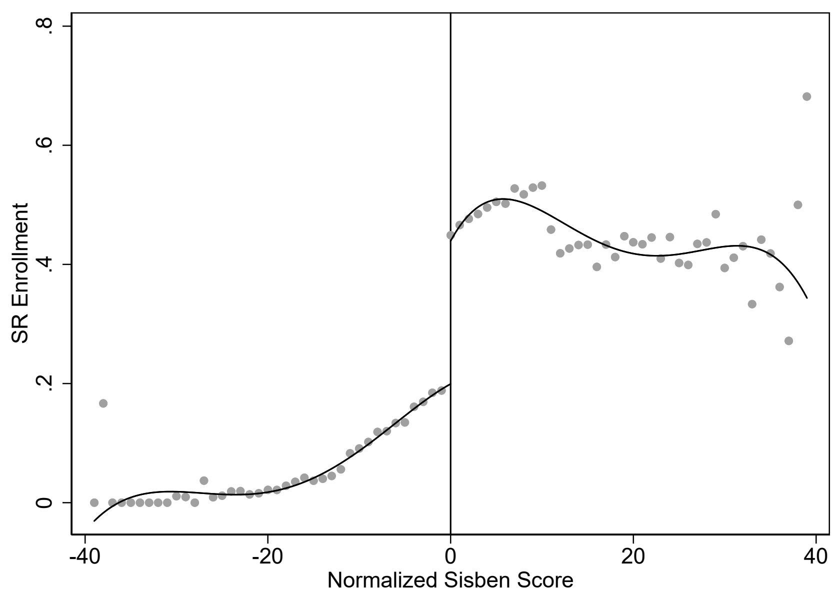
[33]Additionally, using the crime-level classification (rather than the individual flags) of LACE crimes protects us against any police biases against specific individuals, or their characteristics (such as insurance status or who the police have more intelligence on).

4 Enrollment in the Subsidized Regime (SR)

As only households in the two lowest levels of *Sisben* I (2002), a score below 47, could qualify for the SR, we compare households on either side of the cutoff to identify the effect of SR eligibility. First, we verify that there is a discontinuity in the probability of SR enrollment at the cutoff. Second, we examine how the likelihood of being in the formal sector changes at the cutoff. Last, we examine the effect on different types of criminal activity.

Following RD convention, we normalize the *Sisben* score so that treated units are individuals with positive values of our new score. Figure 1 presents the first stage: the discontinuity in the probability of SR enrollment using the optimal binning procedure found in Calonico et al. (2014a). The probability of enrollment discontinuously increases by around 26 percentage points.[34] Not all eligible persons enroll in SR, as formal sector jobs may be valuable to some, but enrollment still jumps substantially to 42% at the cutoff.

Figure 1: Discontinuity in the Probability of SR enrollment at Cutoff.



SR enrollment is probability of being enrolled in the subsidized regime in 2005. RD Graph using optimal binning procedure discussed in Calonico et al. (2014a). Normalized Sisben (2002) score on horizontal axis centered around cutoff. Higher values represent low scores (higher poverty).

For two-staged least squares (2SLS) exercises we follow a fuzzy regression discontinuity design, where our running variable is the 2002 *Sisben* score. We use both parametric and non-

[34] Around 20% of households that have a high 2002 *Sisben* also avail of SR in 2005, as a fraction of households became eligible under a smaller 1998 *Sisben* survey, and the government allows them to keep their benefits for some time after they graduate out of eligibility.

parametric approaches to estimate the effect of SR eligibility at the cutoff. For the parametric approach we follow Hahn et al. (2001), where we instrument enrollment in the SR with the eligibility indicator $1[s_i < 47]$, and estimate the following equation as our first stage:

$$SR_{i,n} = \alpha + \alpha_1 1[s_{i,n} < 47] + X'_{i,n}\alpha_2 + A_i(s_{i,n})\alpha_3 + \mu_n + \varepsilon_{i,n} \ , \tag{1}$$

where A_i is a vector of smooth polynomial functions of the *Sisben* score of each individual, $s_{i,n}$. In robustness checks, we also estimate models conditioning on demographics and other baseline characteristics. Here $X_{i,n}$ is a vector of demographic characteristics for individual i living in neighborhood n. μ_n corresponds to neighborhood fixed effects for the 249 neighborhoods.[35]

An important issue in practice is the selection of the smoothing parameter. We use local regressions to estimate the discontinuity in outcomes at the cutoff point. In particular, we estimate local polynomial regressions conducted with a rectangular kernel and employing the optimal data-driven procedure suggested by Calonico et al. (2014b). We use two different optimal bandwidth procedures: the Imbens and Kalyanaraman (2012) method and the Calonico et al. (2014b) bandwidth. The optimal bandwidths from the different procedures lie between 5.5 and 6.2 points, on the 100-point *Sisben* I scale. We present our results for multiple bandwidths to highlight the robust nature of our estimates, varying them from below the optimal bandwidths to larger bandwidths. Specifically, we check for coefficient stability for results spanning these bandwidths ranging between 4 and 10 points around the cutoff. Varying the size of the bandwidth and the polynomial order do not affect the results.

Our first stage results are shown in Table 2, confirming the 26 percentage point increase in SR enrollment shown in Figure 1. As we vary the bandwidths from 4 through 10 the coefficient is stable and both economically and statistically significant. The table also shows that the standard IV F-test suggests a strong instrument, and for our remaining outcomes we conduct two-staged least squares analyses using this is as our first stage.

[35] We include controls in robustness checks, where we control for various characteristics of the household head in 2002, the baseline year. These controls include an indicator for female-headed households, employment status, years of education, marital status, attendance to any academic institute, year-of-birth fixed effects, socioeconomic strata of the household, and home ownership.

Table 2: SR Enrollment at Sisben Cutoff (First Stage)

Variables	Bandwidths:	4	6	10
	Dependent Variable: Enrolled in SR (First Stage)			
Below Sisben Cutoff		0.260***	0.260***	0.269***
		(0.0138)	(0.0132)	(0.0110)
F-stat of IV		354.97	387.97	598.02
Number of observations		181,132	246,974	340,581
Sample mean (in bandwidth)				0.36

Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. Coefficient of indicator of being below *Sisben* cutoff, with linear controls for 2002 *Sisben* scores that vary flexibly at the cutoff. SR enrollment as measured in the 2005 *Sisben* survey. Standard errors clustered at the *comuna* level.

5 Impacts on Formal Employment and Reported Income

We test the simple hypothesis that the SR conditions disincentivized formal-sector employment and led to an increase in organized-crime activities. We first reproduce a well-established result and show that the program has a negative effect on formal employment (Camacho et al., 2014; Gaviria et al., 2007; Joumard and Londono, 2013; Santamaria et al., 2008). We exploit the discontinuity in enrollment rates at the cutoff, by using the eligibility indicator as an instrument for enrollment status to identify the effect of SR on formal employment and income. Here $Emp_{i,n}$ is 1 if the individual i from neighborhood n was formally employed. We include demographic controls in $X_{i,n}$, and neighborhood fixed effects μ_n. We show the reduced form relationship between employment and being above the RD cutoff:

$$Emp_{i,n} = \gamma_0 + \gamma_1 1\left[s_{i,n} < 47\right] + X'_{i,n}\gamma_2 + A_i\left(s_{i,n}\right)\gamma_3 + \mu_n + \varepsilon_{i,n} \ ,$$

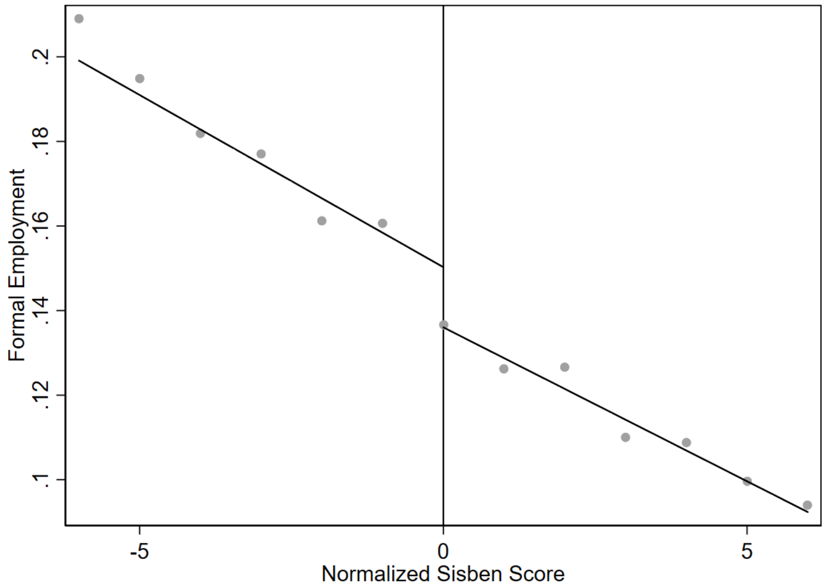
We then instrument for SR enrollment, where $S\hat{R}_{i,n}$ is the predicted SR enrollment probability from the first stage estimated in equation 1. The second stage is:

$$Emp_{i,n} = \beta_0 + \beta_1 S\hat{R}_{i,n} + X'_{i,n}\beta_2 + A_i\left(s_{i,n}\right)\beta_3 + \mu_n + \epsilon_{i,n} \ ,$$

Figure 2 captures the fall in formal sector employment at the cutoff, where formal employment is defined as a working individual making wage contributions to benefits as measured in the

2009 *Sisben* III survey.[36] In all remaining RD figures, we focus on a bandwidth of 6 around the cutoff as it is the Calonico et al. (2014b) optimal bandwidth.

Figure 2: Discontinuity in Formal Employment (2009).



RD Graph using optimal binning procedure discussed in Calonico et al. (2014a). Formal employment based on measures in 2009 *Sisben* survey. Subsample of males. Normalized Sisben (2002) score on horizontal axis centered around cutoff. Higher values represent low scores (higher poverty).

Table 3 presents the results for reported formal employment and incomes in the *Sisben* survey. The table presents results for the reduced form change at the cutoff, and the two-staged least squares (2SLS) effect of enrolling in SR. These results show that the health insurance program had a negative impact of 4.1 percentage points (when using the optimal bandwidth) on the probability of being employed in the formal sector.

Lower formal sector employment at the cutoff may be a combination of fewer youth joining the formal sector as they enter working-age, lower transition rates out of informal work, and higher transition probabilities out of formal work at the cutoff. As formal sector employment affects SR enrollment for the entire family, these are often family decisions, where older family members may discourage youth from joining the formal sector (Joumard and Londono, 2013). This effect is larger for men than it is for women (Appendix Table A2), perhaps once again highlighting that males have an outside option in organized crime.[37]

[36] While this is a somewhat conservative measure of formal employment, paying contributions to health insurance is widely used as a measure of formal employment in Colombia (See Attanasio et al., 2017; Morales and Medina, 2017). The Sisben does not explicitly ask households whether members are in the formal sector.

[37] Note, that we should not necessarily think of this result as a 'first-stage' on crime outcomes. Instead, crime

Table 3: Reported Formal Employment and Income

Bandwidths:	4	6	10
Panel A: Formal Employment in 2009 (Males)			
Above Cutoff	-0.0147***	-0.0111***	-0.00845***
Reduced Form	(0.00467)	(0.00280)	(0.00217)
Enrolled in SR	-0.0539***	-0.0411***	-0.0301***
2SLS	(0.0166)	(0.0103)	(0.00811)
Number of observations	133,067	180,742	247,886
Sample mean (only males in bandwidth for 2009)			0.14
Panel B: Annual Household Income in 2009 (USD)			
Above Cutoff	-3.837	-3.805	2.347
Reduced Form	(3.100)	(2.295)	(4.008)
Enrolled in SR	-6.481	-3.042	30.49
2SLS	(9.163)	(8.842)	(27.83)
Number of observations	46,797	63,457	87,510
Sample mean (households in bandwidth for 2009)			171.24

Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. We use the *Sisben* survey of 2009 to construct both outcome variables. Formal employment for males only. The results for women are presented in Table A2. Tables report Two-Staged Least Squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff. Regressions control linearly for the Sisben score, flexibly around the cutoff. We cluster standard errors by *comuna*. Household-level income reported in *pesos* and converted to USD using the average 2009 exchange rate. Sample means for males and households only in bandwidth for 2009.

The impact on household-level income is statistically indistinguishable from zero and economically small ($30 per household annually). Though we are wary of reading too much into impacts on self-reported income measures, these results suggest that even as workers drop out of the formal sector they find replacement sources of income. One caveat is that income is self-reported, and respondents may under-report assets and incomes in order to get a lower *Sisben* score. However, as respondents do not know the score formula, perfect manipulation is impossible, and so, as we show below, the density of respondents is smooth around the cutoff.

Note, if anything, we may particularly expect that incomes from illicit activities would

and formal employment choices are jointly determined. Indeed, it is possible that, as we predict, the incentives lead individuals to leave the formal sector and join crime. After a few years in crime, an individual may wish to re-join the formal sector, but may be unable to do so given a criminal record.

be under-reported rather than over-reported. Accordingly, the absence of a negative impact on income might suggest that desperation is less likely to be driving any increase in criminal activity we document below. Indeed, by revealed preference, they *choose* to be outside of the formal sector, and as such, should be better off. Nevertheless, we do not wish to emphasize the interpretation of these self-reported income results.

Canonical crime models (Becker, 1968; Ehrlich, 1973) stress the role of both income and substitution effects when wages in one sector change. In other analyses that focus on legitimate sector job-loss and unemployment shocks, the income and substitution effects work to both increase criminal activity. Interestingly, in contrast, here any gains from the subsidy essentially lower the likelihood of criminal activity. The results of this section confirm that the program encouraged informality among young men. The obvious question that this raises is how such discouragement of formal sector employment affects the likelihood of criminal activity.

6 Impacts on Crime

We turn our attention to outcomes on crime. One important distinction with the formal employment results is that we measure crime cumulatively over a decade. We interpret the impacts on crime as causally related to the incentives to leave the formal sector.[38] We show both the reduced form and 2SLS estimates of impacts on crime. In the second stage, we use the eligibility indicator as an instrument for enrollment status to identify the effect of SR enrollment on crime. Here $crime_i$ is 1 if the individual i was arrested between 2005 and 2013.[39]

$$Crime_{i,n} = \beta_0 + \beta_1 S\hat{R}_{i,n} + X'_{i,n}\beta_2 + A_i\left(s_{i,n}\right)\beta_3 + \mu_n + \varepsilon_{i,n}\ ,$$

Our main results do not condition on other factors. In robustness checks, we control for various characteristics of the household head in 2002, the baseline year. These controls include an indicator for female-headed households, employment status, years of education, marital status,

[38]Note that by the latter half of this period almost everyone had healthcare (under either one of the two regimes), and the benefits were similar. As such health benefits are not changing at the cutoff, only the incentives behind who pays for it changes.

[39]Even as we have crime data for many years, we have formal employment only recorded at one point of time in 2009. This poses challenges when trying to simultaneously measure changes in employment and crime. Yet our results are robust to doing so (Appendix Table A9).

attendance to any academic institution, year-of-birth fixed effects, socioeconomic strata of the household,[40] home ownership, and neighborhood fixed effects. A literature on neighborhood effects and crime (Cullen et al., 2006; Dustmann and Damm, 2014) highlight the perils of using area-based relationships (like differences in unemployment rates) to study individual-level occupational choice, and re-iterates the strength of our approach.[41] Our results are unaffected by the inclusion of neighborhood fixed effects that absorb any neighborhood level characteristics (demographics, amenities, property values and police presence) that may affect crime rates. We cluster standard errors at the *comuna* level, but everywhere our results are robust to clustering at other geographic levels, like smaller neighborhoods (*barrios*).

We present results for violent, property, and drug-related crimes, dividing each group between crimes "likely associated with criminal enterprises" (LACE) and those more reflective impulse or opportunity (non-LACE). We hypothesize that LACE crimes are reflective of an occupational choice across legitimate and illegitimate sectors, whereas non-LACE crimes should be less affected by incentives for remaining informal and hence serve as an effective falsification test. We expect the effects on the latter group to be zero, as crimes of impulse and passion are less directly related to occupational choice.

As we elaborate in a later discussion, over and above a falsification test, the lack of effects on non-LACE crimes also allows us to rule out alternative mechanisms. We do not classify crimes based on whether or not they are pecuniary, as that would capture crimes of desperation and necessity that arise out of poverty. Instead, we posit that the policy induced an *occupational choice* into crime, and as such use activities associated with criminal enterprises as a basis for classification. Alternative mechanisms (such as riskier behavior when having insurance) may have weight if non-LACE crimes rose as well, but the lack of effects on non-LACE crimes allows us to rule them out.

We employ conservative outcome definitions: when looking at the impacts on violent crime, we exclude those whose first arrests were in property or drug crime, and do the same for each type of crime. Accordingly, the number of observations differs by type of crime. As such, our

[40] Urban areas in Colombia are split into six socioeconomic strata, used by authorities to spatially target social spending to neighborhoods.

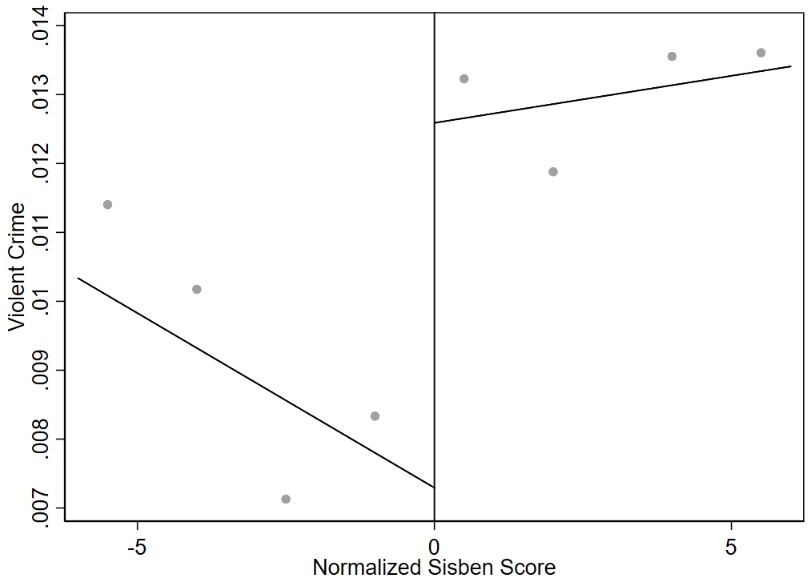
[41] There may still be general equilibrium effects of the policy that affect the entire country, but since our variation is not driven by differences across neighborhoods, unlike other studies, this is all netted out.

outcome will be 1 if the person's first arrest was in violent crime, and 0 if they were never arrested in their youth. In robustness checks, we include the other types of crimes as 0s, and our results are simple more precisely estimated (see Appendix Table A7).

6.1 Violent Crime

We first start with the probability of being arrested for violent criminal activities. Based on the police flags for gang-related activity, violent LACE crimes include homicides, extortion, kidnapping, and firearms trafficking. Non-LACE violent crimes include domestic violence and rape. Figure 3 and Table 4 present the results.

Figure 3: LACE Violent Crimes



RD Graph using optimal binning procedure discussed in Calonico et al. (2014a). Normalized Sisben (2002) score on horizontal axis centered around cutoff. Higher values represent low scores (higher poverty). LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations.

Figure 3 shows the jump in arrests for violent LACE crimes at the *Sisben* cutoff, concentrating on an optimal bandwidth of 6 points on the 100 point scale. In Table 4 we present the regression discontinuity results varying the bandwidth and specifications. The reduced form results (first row in each panel) show an increase in LACE violent crime (Panel A), but no corresponding change in non-LACE violent crime (Panel B). Within a bandwidth of 10 points on the *Sisben* scale, and measuring arrests over a decade, these results amount to a 32% increase (or a 0.45 percentage point increase) in violent crime arrests from the mean around the cutoff.

Table 4: Violent Crimes

Bandwidths:	4	6	10
Panel A: LACE Violent Crimes			
Above Cutoff	0.00722***	0.00649**	0.00456**
Reduced Form	(0.00236)	(0.00249)	(0.00164)
Enrolled in SR	0.0257***	0.0231***	0.0158***
No Covariates	(0.00873)	(0.00838)	(0.00539)
Enrolled in SR	0.0274***	0.0232**	0.0149**
Including pre-treatment covariates	(0.00950)	(0.00937)	(0.00583)
Number of observations	18,052	24,272	33,027
Sample mean (men 13-26 years old in bandwidth)			0.014
Sample mean for those enrolled in SR and in high-gang comuna			0.020
Panel B: Non-LACE Violent Crimes			
Above Cutoff	0.00279	0.000988	-0.000581
Reduced Form	(0.00454)	(0.00304)	(0.00326)
Enrolled in SR	0.00994	0.00349	-0.00201
No Covariates	(0.0158)	(0.0104)	(0.0110)
Enrolled in SR	0.00791	0.00118	-0.00322
Including pre-treatment covariates	(0.0168)	(0.0111)	(0.0125)
Number of observations	18,419	24,768	33,702
Sample mean (men 13-26 years old in bandwidth)			0.034
Sample mean for those enrolled in SR and in high-gang comuna			0.039

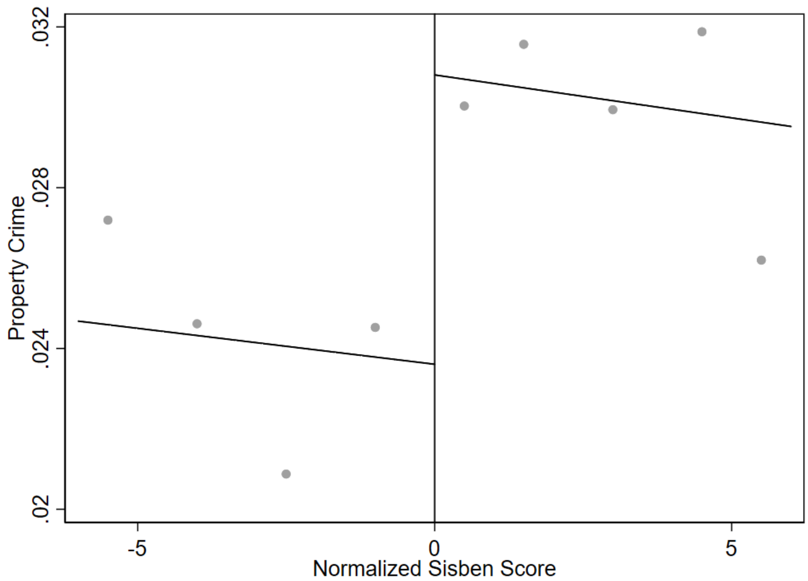
Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Non-LACE crimes are the remaining, more likely representing crimes of impulse or opportunity. Tables report reduced form and two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff. The Sisben score is measured in 2002, and SR enrollment in 2005. We measure crime between 2005 and 2013. Regressions control linearly for the Sisben score, flexibly around the cutoff. We cluster standard errors by *comuna*. We consider only males between 21 to 26 years old in 2013. For regressions that have pre-treatment covariates, we include household characteristics, year of birth fixed effects, and neighborhood fixed effects. The sample excludes anybody whose first arrest was a property or drug crime (Appendix Table A7 includes these observations as a robustness).

These magnitudes are both economically meaningful and similar to those from recent studies in this context (Khanna et al., 2020).

Our 2SLS results (next two rows of each panel) show an economically and statistically significant increase in the probability of arrest for LACE violent crimes for individuals enrolled in SR. We do not find any meaningful effect on the probability of arrest for non-LACE violent crimes. A comparison of the various rows in each panel shows that the estimates are robust to including controls, whereas a comparison across columns shows the robustness to bandwidths. The difference between coefficients for LACE and non-LACE crimes are statistically significantly different for all bandwidths.

6.2 Property Crime

Figure 4: LACE Property Crimes



RD Graph using optimal binning procedure discussed in Calonico et al. (2014a). Normalized Sisben (2002) score on horizontal axis centered around cutoff. Higher values represent low scores (higher poverty). LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations.

In Figure 4 and Table 5 we analyze the effects on property crimes. LACE property crimes include motor vehicle theft and burglary of businesses and residences. Crimes like fraud and identify theft are classified as non-LACE. Once again, we see that LACE property crimes increase, with little change to non-LACE property crimes. The reduced-form estimate, over the entire decade, constitutes a 21% increase (or a 0.66 percentage point increase) from the mean around the cutoff within a bandwidth of 10 points. In 2SLS results, we also find an economically and statistically significant increase for LACE property crime arrests, and no strong effect for property crimes less associated with organized entities. Again, our estimates are quite robust to the inclusion of controls and the choice of the bandwidth. It is interesting

Table 5: Property Crimes

Bandwidths:	4	6	10
Panel A: LACE Property Crimes			
Above Cutoff	0.0106**	0.00930**	0.00666*
Reduced Form	(0.00387)	(0.00389)	(0.00350)
Enrolled in SR	0.0380***	0.0331***	0.0232**
No Covariates	(0.0123)	(0.0126)	(0.0113)
Enrolled in SR	0.0408***	0.0341***	0.0240**
Including pre-treatment covariates	(0.0139)	(0.0131)	(0.0108)
Number of observations	18,426	24,740	33,625
Sample mean (men 13-26 years old in bandwidth)			0.032
Sample mean for those enrolled in SR and in high-gang comuna			0.040
Panel B: Non-LACE Property Crimes			
Above Cutoff	-0.00263	-0.00217	-0.00205
Reduced Form	(0.00554)	(0.00425)	(0.00336)
Enrolled in SR	-0.00941	-0.00772	-0.00712
No Covariates	(0.0194)	(0.0149)	(0.0112)
Enrolled in SR	-0.0116	-0.00872	-0.00854
Including pre-treatment covariates	(0.0212)	(0.0156)	(0.0119)
Number of observations	18,240	24,523	33,358
Sample mean (men 13-26 years old in bandwidth)			0.024
Sample mean for those enrolled in SR and in high-gang comuna			0.028

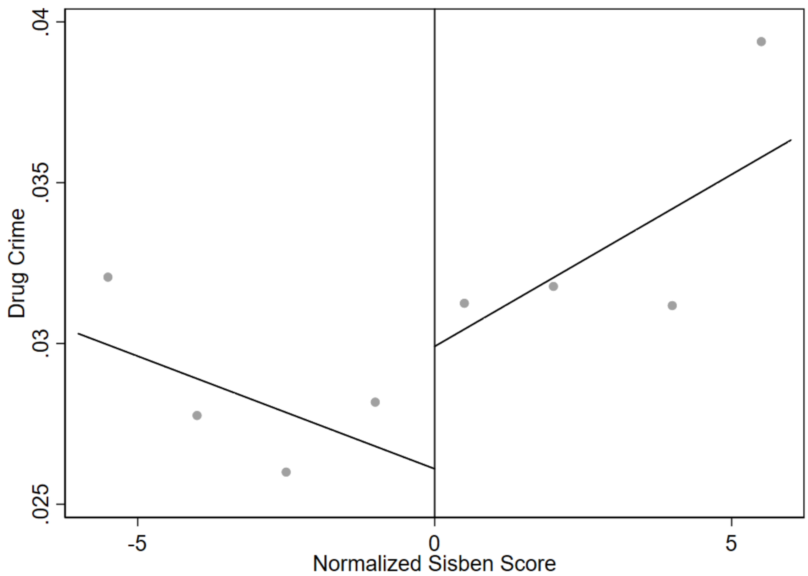
Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Non-LACE crimes are the remaining, more likely representing crimes of impulse or opportunity. Tables report reduced form and two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff. The Sisben score is measured in 2002, and SR enrollment in 2005. We measure crime between 2005 and 2013. Regressions control linearly for the Sisben score, flexibly around the cutoff. We cluster standard errors by *comuna*. We consider only males between 21 to 26 years old in 2013. For regressions that have pre-treatment covariates, we include household characteristics, year of birth fixed effects, and neighborhood fixed effects.The sample excludes anybody whose first arrest was a violent or drug crime (Appendix Table A7 includes these observations as a robustness).

to note that many non-LACE property crimes may also be income generating (even if they do not reflect occupational choices), and as such, impacts on these non-LACE crimes may be consistent with early economic models of crime (Becker, 1968). Yet, we find that it is the decision to join a criminal enterprise that seems to be the driving force. This is consistent with the anthropological interviews discussed above, which illustrate how gangs recruit idle youth and document that working for a gang is lucrative. The difference in coefficients between LACE and non-LACE crimes are statistically significantly different for all bandwidths.

6.3 Drug Crime

We analyze the impact on the probability to engage in drug-related crimes in Figure 5 and Table 6. LACE drug crimes include the manufacturing, distribution, and trafficking of hard drugs like cocaine and heroin. Non-LACE drug crimes include possession and consumption of marijuana, as these are mostly indicative of personal recreational use.

Figure 5: LACE Drug Crimes



RD Graph using optimal binning procedure discussed in Calonico et al. (2014a). Normalized Sisben (2002) score on horizontal axis centered around cutoff. Higher values represent low scores (higher poverty). LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations.

In Figure 5, even though the discontinuity in drug crime arrests is minor, there is a change in the slope of the relationship. The binned averages suggest a somewhat imprecise effect at the cutoff. In Table 6 the direction of effects are what we may expect, but our results are

Table 6: Drug Crimes

Bandwidths:	4	6	10
Panel A: LACE Drug Crimes			
Above Cutoff	0.00799	0.00348	0.00133
Reduced Form	(0.00721)	(0.00492)	(0.00458)
Enrolled in SR	0.0285	0.0124	0.00461
No Covariates	(0.0240)	(0.0169)	(0.0155)
Enrolled in SR	0.0303	0.0135	0.00524
Including pre-treatment covariates	(0.0270)	(0.0180)	(0.0159)
Number of observations	18,463	24,857	33,851
Sample mean (men 13-26 years old in bandwidth)			0.038
Sample mean for those enrolled in SR and in high-gang comuna			0.045
Panel B: Non-LACE Drug Crimes			
Above Cutoff	-0.00976	-0.0129	-0.00788
Reduced Form	(0.00774)	(0.00798)	(0.00629)
Enrolled in SR	-0.0348	-0.0458	-0.0274
No Covariates	(0.0280)	(0.0293)	(0.0218)
Enrolled in SR	-0.0385	-0.0501	-0.0277
Including pre-treatment covariates	(0.0299)	(0.0329)	(0.0230)
Number of observations	19,150	25,740	35,104
Sample mean (men 13-26 years old in bandwidth)			0.073
Sample mean for those enrolled in SR and in high-gang comuna			0.088

Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Non-LACE crimes are the remaining, more likely representing crimes of impulse or opportunity. Tables report reduced form and two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff. The Sisben score is measured in 2002, and SR enrollment in 2005. We measure crime between 2005 and 2013. Regressions control linearly for the Sisben score, flexibly around the cutoff. We cluster standard errors by *comuna*. We consider only males between 21 to 26 years old in 2013. For regressions that have pre-treatment covariates, we include household characteristics, year of birth fixed effects, and neighborhood fixed effects. The sample excludes anybody whose first arrest was a property or violent crime (Appendix Table A7 includes these observations as a robustness).

not precisely estimated. In Appendix Table A9, we use an alternative joint outcome definition of arrested and informal in 2009. Those results are precise enough to measure a significant increase in LACE drug crimes at the cutoff.

One possibility for the reduced precision in arrests for drug crimes is measurement error associated with the classification of such crimes. That is, the difficulty in classifying possession of drugs as consumption versus trafficking or distribution likely introduces noise. Indeed, offenses related to the trafficking of marijuana are problematic as small amounts of personal possession were made legal during this period. While homicides, assaults and theft involve victims as clear evidence of crimes, drug crimes are often difficult to detect and record. Not having any evidence of a crime actually being committed (e.g., a victim) may also allow authorities to under-report, especially if cartels pressure authorities to do so.

In sum, our results indicate that the drop in formal employment as a result of the subsidized benefits for informal workers raised the likelihood of being arrested for LACE violent and property crimes. The magnitudes of the estimated impacts are also economically meaningful. The pattern of results is similar but imprecise for drug crimes. Importantly, the results also show that non-LACE crimes of each type are not impacted by SR enrollment, ruling out many alternative mechanisms. In the following section, we investigate whether impacts are strongest in *comunas* that were historically associated with high organized crime activity as further evidence in support of our occupational choice interpretation.

7 Heterogeneity, Specification Tests and Robustness

7.1 Heterogeneity by Comuna: the Importance of Neighborhoods

Previous studies have emphasized that the opportunities in a neighborhood affect how easy it is to induce youth into crime (Kling et al., 2005). Understanding the heterogeneity by neighborhood helps us speak to much of the literature which relies on area-based variation. High crime neighborhoods may have more policing and higher detection rates that may lower the employment-crime elasticity, but may also have more opportunities to join a gang and thereby raise the elasticity.

Table 7: Heterogeneity by Comuna

Bandwidths:	4	6	10
Panel A: LACE Violent Crimes			
Enrolled in SR	0.0267***	0.0211**	0.0150***
	(0.00892)	(0.00914)	(0.00538)
Enrolled* Gang Comuna	-0.00152	0.0141***	0.00563
	(0.00464)	(0.00376)	(0.00537)
F stat	90.2	154.6	232.9
Number of observations	18,052	24,272	33,027
Panel B: LACE Property Crimes			
Enrolled in SR	0.0344**	0.0273**	0.0190*
	(0.0134)	(0.0137)	(0.0115)
Enrolled* Gang Comuna	0.0282	0.0364**	0.0258**
	(0.0209)	(0.0167)	(0.0116)
F stat	86.7	145.5	249.6
Number of observations	18,426	24,740	33,625
Panel C: LACE Drug Crimes			
Enrolled in SR	0.0282	0.0131	0.00296
	(0.0248)	(0.0177)	(0.0166)
Enrolled* Gang Comuna	0.000310	-0.00590	0.00690
	(0.0136)	(0.0135)	(0.0129)
F stat	96	149.1	201.7
Number of observations	18,463	24,857	33,851

Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Tables report two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff and an interaction between high-gang comunas and being below the cutoff. The Sisben score is measure in 2002, SR enrollment in 2005, and crime outcomes are measured between 2005 and 2013. Regressions include comuna fixed effects and an interaction between high-gang comunas and indicators for SR enrollment. Regressions control linearly for the Sisben score, flexibly around the cutoff. We consider only males between 21 to 26 years old in 2013. We cluster errors by comuna. The mean arrest rate across all five gang comunas are 18%, which is also the mean arrest rate in non-gang comunas. See Appendix Table A4 for the non-LACE crimes.

We investigate if *comunas* with a high incidence of gangs demonstrate stronger impacts on LACE arrests at the RD cutoff. If the policy induces men to join organized crime, then we may expect that neighborhoods that have more such opportunities would have a larger impact. Figure A3 shows the spatial distribution of the locations where criminals were arrested in the act between 2005 and 2013, by type of crime.[42]

We select the five *comunas* with the highest number of gang members captured by the police, and create an indicator variable for whether individuals lived in these *comunas* in 2002, our baseline year.[43] These are not necessarily high crime areas, as the mean arrest rates for young men is 18% in both gang and non-gang *comunas*. Yet, LACE crimes make up 43% of arrests in gang *comunas*, and 37% of arrests in non-gang *comunas* .

We interact this variable with the cutoff to analyze the heterogeneity in effects by area-level gang activity. Table 7 presents the results. Since we have an interaction term, we report the IV first stage F-statistics as well. The effects on crime are present in both high and low gang activity areas, but for property crime are larger in areas that have more gang activity. For violent crime the interaction term is strongly positive for the optimal bandwidth, but less robust than results for property crimes.

Appendix Table A4 shows the results for non-LACE crimes. Again, there is no evidence of SR enrollment being associated with non-LACE crimes in either gang *comunas* or non-gang *comunas*. Notice that our identification strategy protects against increases in policing activity in gang comunas, as we are comparing one side of the *Sisben* cutoff to the other. Additionally, throughout the paper our tables show a row of results that include *comunas* fixed effects.

7.2 Density Tests and Balance Tests

In our study, identification relies on the assumption that all other determinants of the outcome vary smoothly at the cutoff. The SR *Sisben* cutoff did not coincide with other interventions, at the eligibility cutoff of *Sisben* during the study period in Medellín. Importantly, we show that an extensive set of observables display no systematic patterns in discontinuities. In Table A3,

[42] In our main results we already show specifications that include neighborhood fixed effects, and we cluster errors at spatial levels larger than neighborhoods. Our results are robust to clustering at smaller spatial levels, like the neighborhood.

[43] The top five are chosen on the criteria of having the most gang-flags as a ratio of total crimes.

we show that baseline characteristics from 2002 (three years before our crime data begins) are balanced for the entire sample and for the sample used in the regressions (young males), respectively. We consider two sets of baseline characteristics: one for household-level socioeconomic variables, and the other for individuals. We find no evidence of systematic discontinuities in covariates at the threshold. In the first row of Table A3, we report a summary measure in which we collapse these variables by taking their first principal component and repeat the same RD analysis that we do for our main results. Again, there is no detectable difference in this composite measure of baseline characteristics. In all our regression tables we show effects both with and without these variables as controls.

Additionally, for the empirical strategy to be valid, households must not be able to manipulate their score to cross the cutoff. Work by Camacho and Conover (2011) highlights politically motivated manipulation in 1998 in other parts of Colombia during mayoral elections. This includes both under-reporting of wealth (not necessarily a threat to our design), but also manipulation of the score. We use the raw survey data and the 2002 *Sisben* score only for Medellín, and so are not concerned with any manipulation of the final score. Indeed, our tests of balance in the large set of baseline characteristics of the household are indicative of a lack of systematic manipulation in this context. Importantly, the cutoff for SR eligibility was determined well after the 2002 *Sisben* scores were released, and as such the cutoffs were not known to anybody during the survey. As such, we are confident that nobody could manipulate the 2002 score, even if they joined a life of crime in later years.

We test whether there was a discontinuity in the density of scores at the cutoff for the particular context of Medellín after 2002. We do this by following two methods used in the literature: the McCrary (2008) test and a test recently developed by Cattaneo et al. (2017). The Cattaneo et al. (2017) test yields a conventional t-statistic of 0.0489 or a p-value of 0.961, and a robust bias corrected p-value of 0.940, confirming that there is no statistically detectable evidence of manipulation. Figure A4 shows the distribution of the *Sisben* score for males (non-criminals and criminals) and for the full sample, conducts McCrary (2008) tests, and shows a closeup of the distribution around the cutoff. The distribution is smooth with no evidence of bunching before the cutoff between *Sisben* levels 2 and 3 (red line).

7.3 Alternative Crime Classifications, Different Bandwidths, Specifications and Subsamples, and the Nonparametric RD

We conduct a number of robustness checks. First, we re-classify crimes into LACE and non-LACE groups based on the *location* where these types of crimes are more likely to occur. We calculate the relative propensity of each crime in each neighborhood. The crimes that have a higher propensity to take place in neighborhoods that were traditionally associated with organized crime are classified as LACE crimes. These are neighborhoods that also have the highest proportion of gang flags associated with them. To be specific, we sort the crimes by the fraction of first arrests that happen in a high-gang neighborhood. The top half of this list is classified as LACE crimes.

This 'Neighborhood Classification Method' of crimes produces a list similar to the one where we use the police generated flags, with minor differences.[44] The lists of the top prevalent crimes by classification method can be found in Appendix Table A1. In Table A5, we re-examine our main results using the alternative classification for LACE crimes. These results are similar, but with the added statistical significance of drug crimes under some specifications.

Next, we re-examine our main results using the bias-correction methods suggested by Calonico et al. (2014a). In Table A6, we show results that conduct a polynomial bias correction at a larger bias-correction bandwidth (reported in the table). Once again, our results show an economically and statistically significant increase in LACE violent and property crimes, but the effects on drug crimes are small and imprecise.

In our main specifications, when looking at a specific type of crime, we exclude arrests from other crimes.[45] In the specification shown in Table A7 we include the other categories along with the non-criminals, and in Table A8 we include repeat arrests. Our results are similarly robust. In Figure A5, we vary the bandwidth through a much wider range – every integer between 2 and 10. LACE violent and property crimes consistently display a positive RD coefficient, whereas drug crimes are not statistically indistinguishable from zero, even as the coefficients

[44]The big differences among the top 25 crimes is with 'conspiracy to commit murder' that is classified as LACE under the original classification but not LACE under the neighborhood definition. Outside the top 25 there are other differences in the definition.

[45]For instance, when studying violent crimes, we exclude property and drug crimes from the sample altogether.

are positive and fairly large for smaller bandwidths.

Finally, as our hypothesis is about both informality and violent crime, we present a specification in Table A9 that simultaneously captures both. The advantage of our individual-level data is that we can measure both employment and criminal behavior for the same individual. However, since we do not have annual data on formal employment, we use the 2009 *Sisben* to measure formal employment. Here the dependent variable is an indicator that equals one when the individual was not formally employed and arrested for a crime, and zero otherwise. Our results again show an increase in LACE criminal activity, with even the drug crimes now being economically and statistically significant. This result allows us to address any concerns that the increase in informality and increase in arrests were coincidental.

8 Interpretation and Alternative Mechanisms

The simultaneous decrease in formal sector employment and rise in LACE related arrests supports a model in which discouragement from formal employment induces an occupational choice into a life of crime (Becker, 1968).[46] Indeed, the results for joint informality and LACE arrest outcomes, presented in Table A9 and discussed above, validate the interpretation of our results as being driven by individual-level occupational choice. Given the lack of effects for non-LACE crimes, it is difficult to find alternative explanations to reconcile these results.

We consider three alternative theories. First, better health benefits at the cutoff may induce one to engage in riskier behaviors. However, it is difficult to support why these riskier behaviors would not also include non-LACE crimes (e.g., drug consumption). Furthermore, as health coverage is near universal by the end of this period, individuals on both sides of the cutoff have nearly identical coverage, with the only difference being who pays.[47] This suggests that the driving force is the cost of coverage under formal sector employment.

Second, formal workers vesting more into the health system may fear losing their jobs if arrested and reduce criminal activities as a result. However, again this should be just as true

[46] While we do not discuss in detail specific pathways, anthropological evidence lends credence to active recruitment by gang members of young men that 'hang around' in neighborhoods with idle time, and are not in the formal sector (Baird, 2011).

[47] If anything, any early differences in coverage, favored better care for formal employees, which would lead to effects of the opposite sign of what we find.

for non-LACE crimes. Third, the police may falsely target informal workers even if they are not criminals, but it is unlikely that they would be booked disproportionately under LACE crimes. Indeed, it may be easier to falsely target potential criminals for petty crime rather than more serious offenses like homicide or auto theft. The distinction between LACE and non-LACE crimes powerfully helps exclude alternative mechanisms and lends credence to the occupational choice story we posit.

We find that a 1 percentage point fall in formal employment is related to 3.1% increase in arrests, which lies in the lower range of recently reviewed estimates.[48] Our estimates reflect slight differences with other estimates in the literature. Becker (1968) discusses both income and substitution effects when wages rise in one sector. For much of the other literature, job loss or unemployment in the formal sector may produce income and substitution effects that both accentuate criminal activity. In contrast, we may think that the gains from the subsidy in our case may actually reduce criminal activity. Additionally, while we are directly testing the choice of occupation, most of the related work studies reduced access to legitimate jobs as a consequence of job losses or work restrictions, and thereby recover different elasticities.[49]

As our estimates are similar to those obtained from other recent causal analysis leveraging individual-level variation (Khanna et al., 2020; Rose, 2019), we interpret our results to be not only economically meaningful, but also plausible. Yet, these magnitudes should be understood to be context-specific. We study a high-crime environment, similar only to other developing countries and especially Latin America.[50] Furthermore, we estimate a Local Average Treatment Effect (LATE) on marginal workers in the neighborhood of an income cutoff. It is plausible that at higher income levels, healthcare is a less important fraction of expenditures, and is less likely to induce such behavior. Finally, it should be noted that the newly induced marginal criminals may be unlike the average criminal along many dimensions, including ease of avoiding arrests, and as such our results may not be widely generalizable for other sub-populations. Nevertheless,

[48] A recent review by Bennett and Ouazad (2018) discusses how a 1 percentage point increase in unemployment usually corresponds to a 3-7% increase in crime. In other work on Medellín, we show that for mass-layoffs among the adult population across all genders, a 1 percentage point increase in unemployment raises arrests by 2.12% (Khanna et al., 2020).

[49] Job losses and structurally imposed employment restrictions may additionally induce effects on depression, subsequent job search, social stigma etc. whereas our variation in relative costs of employment should not.

[50] Recent evidence from Latin America suggests much larger employment-crime elasticities (Dell et al., 2018; Dix-Carneiro et al., 2018).

since the exogenous probability of getting caught conditional on committing a crime has no reason to be discontinuous at the cutoff, our estimates are unbiased even in the presence of such heterogeneity in criminal "skill."

9 Conclusion

In this paper, we highlight an important fact: disincentivizing formal employment can lead to substantial increases in criminal activity when informal opportunities include employment by criminal enterprises. We evaluate this claim in the context of the high-crime environment of Medellín, Colombia. We first provide strong evidence showing that the criteria behind the health benefits policy led to a sharp decrease in formal sector employment. At the margin, the policy encouraged workers to remain in the informal labor market.

In Medellín, this informal market contains significant opportunities related to organized crime. We follow these same individuals over a decade and show that this decrease in formal sector employment led to an increase in the probability of being arrested for criminal-enterprise related activities. On the other hand, crimes less likely to be associated with criminal economic enterprises, like crimes of impulse or opportunity, show no such impacts at the eligibility threshold, lending credence to the occupational choice mechanism we advance. Together, our simple calculations suggest that as the policy pushed workers out of the formal sector, a meaningful fraction of these workers were drawn into criminal enterprises. These effects were largest in neighborhoods that had, at baseline, greater opportunities to join organized crime.

Crime deterrence may have limited benefits if the supply elasticity to criminal activity is high (Freeman, 1999). Investigating the decisions behind choosing a life of crime, as we do here, is essential in the fight against crime. Importantly, our work speaks to the determinants of engaging in criminal activity at the individual level. The strength of our approach is that we do not need to use area-based variation or disequilibrium shocks like job-losses to identify the occupation choice between individual employment opportunities and crime. We do this using a unique data set that matches the census of arrests with socio-economic outcomes over a decade, in the context of one of the most violent cities in the world. We find a source of exogenous variation generated by policy rules, and use an RD design to estimate our effects.

We conclude that Colombia's well-intentioned and broad-based subsidies for healthcare had the unintended consequence of incentivizing participation in criminal enterprise by way of its distortionary provision rules. The program being important for providing subsidized health access to low income families implies that there is little reason to do away with it. Yet, the formality-clause governing the selection into the program is distortionary, and as such warrants examination.[51]

Removing the emphasis on informality (but still targeting the poor) may negate the increase in criminal activity around the cutoff. The costs underlying such a change would be a larger fiscal burden as even low-income formal sector workers would be eligible for SR. The benefits are far reaching: less crime, less policing and incarceration, and fewer negative externalities on families and children. This has welfare implications for the design of many programs across the developing world which often have far-reaching and under-studied consequences on seemingly unrelated outcomes and behaviors. Our results provide guidance for how impactful improving access to and incentives for formal sector employment can be for deterring criminal activity.

[51] Recognizing these adverse effects, policy-makers lowered the costs of CR enrollment at the end of our study period, when Law 1607 was enacted, leading to a significant increase in formal sector employment (Bernal et al., 2017; Kugler et al., 2017; Morales and Medina, 2017).

References

Agan, A. and M. Makowsky (2018). The Minimum Wage, EITC, and Criminal Recidivism. *Working paper*.

Agan, A. and S. Starr (2018). Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment. *The Quarterly Journal of Economics 133*(1), 191–235.

Alesina, A., S. Piccolo, and P. Pinotti (2017). Organized Crime, Violence, and Politics. *NBER Working Paper*.

Arteaga, C. (2019). The Cost of Bad Parents: Evidence from the Effects of Parental Incarceration on Children's Education. *Working Paper*.

Attanasio, O., A. Guarín, C. Medina, and C. Meghir (2017). Vocational Training for Disadvantaged Youth in Colombia: A Long-Term Follow-up. *American Economic Journal: Applied Economics 9*(2), 131–43.

Baird, A. (2011). Negotiating Pathways to Manhood: Violence Reproduction in Medellin's Periphery. *Thesis, University of Bradford*.

Barnwal, P. (2018). Curbing Leakage in Public Programs: Evidence from India's Direct Benefit Transfer Policy. *Mimeo*.

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

Bell, B., A. Bindler, and S. Machin (2018). Crime scars: Recessions and the making of career criminals. *Review of Economics and Statistics 100*(3), 392–404.

Bennett, P. and A. Ouazad (2018). Job Displacement, Unemployment, and Crime: Evidence from Danish Microdata and Reforms. *Journal of the European Economic Association*. forthcoming.

Bernal, R., M. Eslava, M. Melendez, and A. Pinzon (2017). Switching from Payroll Taxes to Corporate Income Taxes: Firms' Employment and Wages after the Colombian 2012 Tax Reform. *Inter-American Development Bank No. 1268*.

Bhuller, M., G. Dahl, K. Loken, and M. Mogstad (2018). Incarceration, Recidivism and Employment. *NBER Working Paper No. 22648*.

BJS, U. B. o. J. S. (1994). Sourcebook of Criminal Justice Statistics. *University at Albany*.

Blattman, C. and J. Annan (2015). Can Employment Reduce Lawlessness and Rebellion? A Field Experiment with High-Risk Men in a Fragile State. *American Political Science Review 10*(1), 1–17.

Blattman, C., G. Duncan, L. Benjamin, and S. Toben (2018). Gangs of Medellín: How Organized Crime is Organized. Technical report.

Brown, R. and A. Velasquez (2017). The Effect of Violent Crime on the Human Capital Accumulation of Young Adults. *Journal of Development Economics 127*, 1–12.

Buonanno, P. and J. Vargas (2018). Inequality, crime, and the long run legacy of slavery. *Journal of Economic Behavior and Organization*. forthcoming.

Calonico, S., M. Cattaneo, and R. Titiunik (2014a). Robust Data-Driven Inference in the Regression Discontinuity Design. *Stata Journal 14*(4), 909–946.

Calonico, S., M. Cattaneo, and R. Titiunik (2014b). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica 82*(6), 2295–2326.

Camacho, A. and E. Conover (2011). Manipulation of Social Program Eligibility. *American Economic Journal: Economic Policy 3*(2), 41–65.

Camacho, A., E. Conover, and A. Hoyos (2014). Effects of Colombia's Social Protection System on Workers' Choice between Formal and Informal Employment. *World Bank Economic Review 28*(3), 446–466.

Cattaneo, M., M. Jansson, and X. Ma (2017). Simple Local Regression Distribution Estimators with an Application to Manipulation Testing. *Mimeo Michigan*.

CCSPJP (2009). Consejo Ciudadano para la Seguridad Publica y la Justicia Penal.

Chimeli, A. B. and R. R. Soares (2017). The use of violence in illegal markets: Evidence from mahogany trade in the brazilian amazon. *American Economic Journal: Applied Economics 9*(4), 30–57.

Chioda, L., J. De Mello, and R. R. Soares (2016). Spillovers from Conditional Cash Transfer Programs: Bolsa Familia and Crime in Urban Brazil. *Economics of Education Review 54*, 306–320.

Cornwell, C. and W. N. Trumbull (1994). Estimating the Economic Model of Crime with Panel Data. *Review of Economics and Statistics 76*(2), 360–366.

Cortes, D., J. Santamaria, and J. Vargas (2016). Economic shocks and crime: Evidence from the crash of ponzi schemes. *Journal of Economic Behavior and Organization 131*, 263–275.

Cullen, J. B., B. A. Jacob, and S. D. Levitt (2005). The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools. *Journal of Public Economics 89*(6), 729–60.

Cullen, J. B., B. A. Jacob, and S. D. Levitt (2006). The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica 74*(5), 1191–1230.

Cullen, J. B. and S. D. Levitt (1999). Crime, Urban Flight, and the Consequences for Cities. *The Review of Economics and Statistics 81*(2), 159–169.

Dal Bó, E. and P. Dal Bó (2011). Workers, warriors, and criminals: social conflict in general equilibrium. *Journal of the European Economic Association 9*(4), 646–677.

Dell, M., B. Feigenberg, and K. Teshima (2018). The Violent Consequences of Trade-Induced Worker Displacement in Mexico. *American Economic Review: Insights forthcoming*.

DellaVigna, S. and D. M. Paserman (2005). Job Search and Impatience. *Journal of Labor Economics 23*(3), 527–588.

DiTella, R., S. Galiani, and E. Schargrodsky (2010). Crime distribution and victim behavior during a crime wave. In *The economics of crime: Lessons for and from Latin America*, pp. 175–204. University of Chicago Press.

Dix-Carneiro, R., R. Soares, and G. Ulyssea (2018). Economic Shocks and Crime: Evidence from the Brazilian Trade Liberalization. *American Economic Journal: Applied Economics forthcoming*.

Doyle, C. (2016). Explaining Patterns of Urban Violence in Medellin, Colombia. *Laws 5*(3).

Doyle, J. (2008). Child Protection and Adult Crime. *Journal of Political Economy 116*(4), 746–770.

Doyle, J. J. (2007). Child Protection and Child Outcomes: Measuring the Effects of Foster Care. *American Economic Review 97*(5), 1583–1610.

Dube, O. and J. F. Vargas (2013). Commodity Price Shocks and Civil Conflict: Evidence from Colombia. *The Review of Economic Studies 80*(4), 1384–1421.

Dustmann, C. and A. P. Damm (2014). Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior? *American Economic Review 104*(6), 1806–1832.

Ehrlich, I. (1973). Participation in Illegitimate Activities: A Theoretical and Empirical Investigation. *Journal of Political Economy 81*(3), 521–565.

Entorf, H. (2000). Socioeconomic and Demographic Factors of Crime in Germany: Evidence from Panel Data of the German States. *International Review of Law and Economics 20*(1), 75–106.

Feler, L. and M. Z. Senses (2017). Trade Shocks and the Provision of Local Public Goods. *American Economic Journal: Economic Policy 9*(4), 101–43.

Fella, G. and G. Gallipoli (2014). Education and Crime over the Life Cycle. *The Review of Economic Studies 81*(4), 1484–1517.

Fernández, C. and L. Villar (2017). The Impact of Lowering the Payroll Tax on Informality in Colombia. *Economía 18*(1), 125–155.

Ferraz, C. and F. Finan (2008). Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. *The Quarterly Journal of Economics 123*(703-745), 2.

Ferraz, C. and F. Finan (2011). Electoral Accountability and Corruption: Evidence from the Audit Reports of Local Governments. *American Economic Review 101*(4), 1274–1311.

Foley, C. F. (2011). Welfare Payments and Crime. *Review of Economics and Statistics 93*(1), 97–112.

Freeman, R. B. (1999). The Economics of Crime. *Handbook of Labor Economics 3*(c). edited by O. Ashenfelter and D. Card. Elsevier Science.

Gathmann, C., I. Helm, and U. Schönberg (2020). Spillover Effects of Mass Layoffs. *Journal of the European Economic Association*. Forthcoming.

Gaviria, A., C. Medina, and C. Mejia (2007). Assessing Health Reform in Colombia: From Theory to Practice. *Economia 7*(1).

Gould, E. D., B. A. Weinberg, and D. B. Mustard (2002). Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997. *The Review of Economics and Statistics 84*(1), 45–61.

Greenbaum, R. T. and G. E. Tita (2004). The Impact of Violence Surges on Neighborhood Business Activity. *Urban Studies 13*, 2495–2514.

Grogger, J. (1995). The Effect of Arrests on the Employment and Earnings of Young Men. *Quarterly Journal of Economics 110*(1), 51–72.

Gronqvist, H. (2017). Youth Unemployment and Crime: Lessons from Longitudinal Population Records. *Mimeo*.

Guarin, A., C. Medina, and J. Tamayo (2013). The Effects of Punishment of Crime in Colombia on Deterrence, Incapacitation and Human Capital Formation. *Borradores de Economia 774*.

Hahn, J., P. Todd, and W. van der Klaauw (2001). Identification and Estimation of Treatment Effects with a Regression Discontinuity Design. *Econometrica 69*(1), 201–209.

Heller, S. (2014). Summer Jobs Reduce Violence among Disadvantaged Youth. *Science 346*(6214), 1219–1223.

Hjalmarsson, R. and M. Lindquist (2019). The Causal Effect of Military Conscription on Crime. *The Economic Journal 129*(622), 2522–2562.

Ibanez, A. M., A. Ritterbusch, and C. Rodriguez (2017). Impact of the Judicial System Reform on Police Behavior: Evidence on Juvenile Crime in Colombia. *Mimeo*.

Ihlanfeldt, K. (2007). Neighborhood Drug Crimes and Young Males Job Accessibility. *The Review of Economics and Statisticsc 89*(1), 151–164.

Imbens, G. W. and K. Kalyanaraman (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies 79*(3), 933–959.

Joumard, I. and J. Londono (2013). Income Inequality and Poverty in Colombia: The Role of the Labour Market. *OECD Economics Department Working Papers No. 1036* (1).

Karin, E. (2005). Unemployment and Crime: Is There a Connection? *Scandinavian Journal of Economics 107*(2), 353–373.

Kearney, M. S., B. Harris, E. Jacome, and L. Parker (2014). Ten Economic Facts about Crime and Incarceration in the United States. *The Hamilton Project Policy Memo*. Brookings Institute.

Khanna, G., C. Medina, A. Nyshadham, C. Posso, and J. Tamayo (2020). Job Loss, Credit and Crime in Colombia. *American Economic Review: Insights*. Forthcoming.

Kling, J. (2006). Incarceration Length, Employment, and Earnings. *American Economic Review*.

Kling, J., J. Ludwig, and L. Katz (2005). Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment. *The Quarterly Journal of Economics 120*(1), 87–130.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*(1), 83–119.

Kugler, A. and M. Kugler (2009). Labor Market Effects of Payroll Taxes in Developing Countries: Evidence from Colombia. *Economic Development and Cultural Change 57*(2), 335–358.

Kugler, A., M. Kugler, and L. Herrera Prada (2017). Do Payroll Tax Breaks Stimulate Formality? Evidence from Colombia's Reform. *NBER Working Paper No. 23308*.

Lamprea, E. and J. Garcia (2016, December). Closing the Gap Between Formal and Material Health Care Coverage in Colombia. *Health and Human Resources Journal*.

Lin, M.-J. (2008). Does Unemployment Increase Crime? Evidence from U.S. Data 1974-2000. *Journal of Human Resources 43*(2), 413–436.

Lochner, L. and E. Moretti (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *American Economic Review 94*(1).

Lott, J. (1992). An Attempt at Measuring the Total Monetary Penalty from Drug Convictions: The Importance of an Individual's Reputation. *Journal of Legal Studies 29*, 159–187.

Machin, S. and C. Meghir (2004). Crime and Economic Incentives. *Journal of Human Resources 39*(4), 958–979.

McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics 142*(2), 698–714.

Medina, C., J. Núñez, and J. Tamayo (2013). The unemployment subsidy program in colombia: An assessment. Technical report.

Medina, C. and J. Tamayo (2011). An Assessment of How Urban Crime and Victimization Affects Life Satisfaction. *in Dave, Webb and Eduardo, Wills-Herrera (Eds.), Subjective Well-Being and Security*. Springer, Social Indicators Research Series.

Melnikov, N., C. Schmidt-Padilla, and M. M. Sviatschi (2019). Gangs, labor mobility, and development. *Working Paper*.

Morales, L. and C. Medina (2017). Assessing the Effect of Payroll Taxes on Formal Employment: The Case of the 2012 Tax Reform in Colombia. *Economia Journal 0*, 75–124. LACEA.

Mustard, D. B. (2010). How Do Labor Markets Affect Crime? New Evidence on an Old Puzzle. *IZA Discussion Paper* (4856).

Olken, B. and R. Pande (2012). Corruption in Developing Countries. *Annual Review of Economics 4*(1), 479–509.

Pinotti, P. (2017). Clicking on Heaven's Door: The Effect of Immigrant Legalization on Crime. *American Economic Review 107*(1), 138–168.

Raphael, S. and R. Winter-Ember (2001). Identifying the Effect of Unemployment on Crime. *Journal of Law and Economics 44*(1), 259–283.

Rose, E. (2019). The Effects of Job Loss on Crime: Evidence From Administrative Data. *Working Paper*.

Rozema, R. (2018). Urban DDR-processes: Paramilitaries and Criminal Networks in Medellin, Colombia. *Journal of Latin American Studies 40*, 423–452.

Santamaria, M., F. Garcia, and A. V. Mujica (2008). Los Costos No Laborales y el Mercado Laboral: Impacto de la Reforma de Salud en Colombia. *Working Paper 43. Fedesarrollo, Bogota, Colombia*.

Schnepel, K. T. (2018). Good jobs and recidivism. *The Economic Journal 128*(608), 447–469.

Schochet, P., J. Burghardt, and S. McConnell (2008). Does job corps work? impact findings from the national job corps study. *American Economic Review 98*(5), 1864–1886.

Sviatschi, M. M. (2018). Making a Narco: Childhood Exposure to Illegal Labor Markets and Criminal Life Paths. *Working paper*.

Sviatschi, M. M. (2020). Spreading Gangs: Exporting US Criminal Capital to El Salvador. *Mimeo*.

Velasquez, A. (2020). The Economics Burden of Crime: Evidence from Mexico. *The Journal of Human Resources*. Forthcoming.

Yang, D. (2008). Can Enforcement Backfire? Crime Displacement in the Context of Customs Reform in the Philippines. *Review of Economics and Statistics 90*(1), 1–14.

Online Appendix: Additional Tables and Figures

Figure A1: Timeline of Data Used

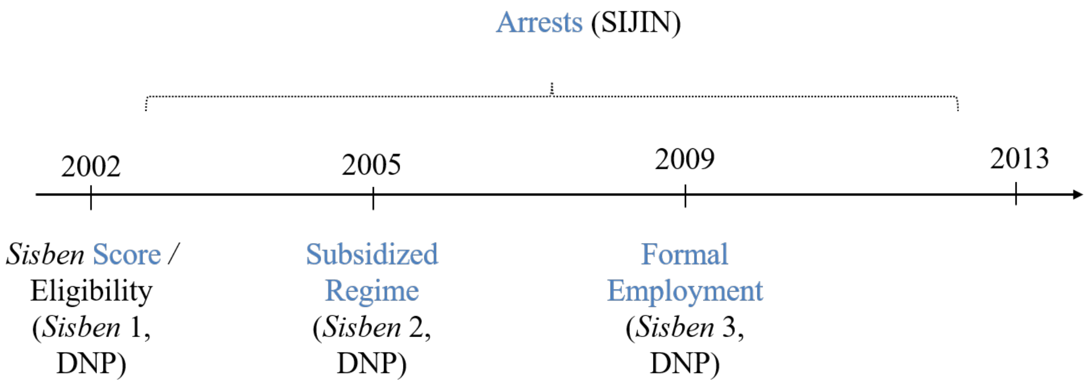
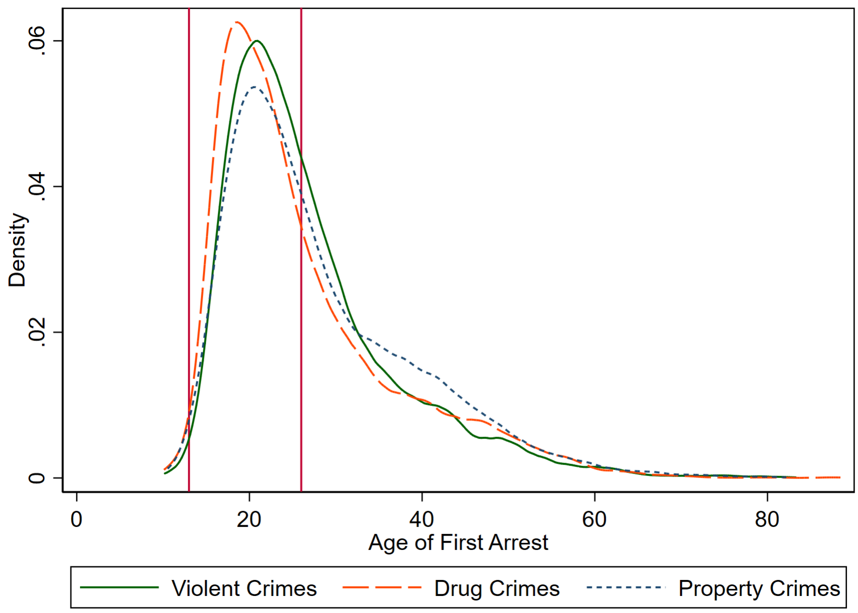
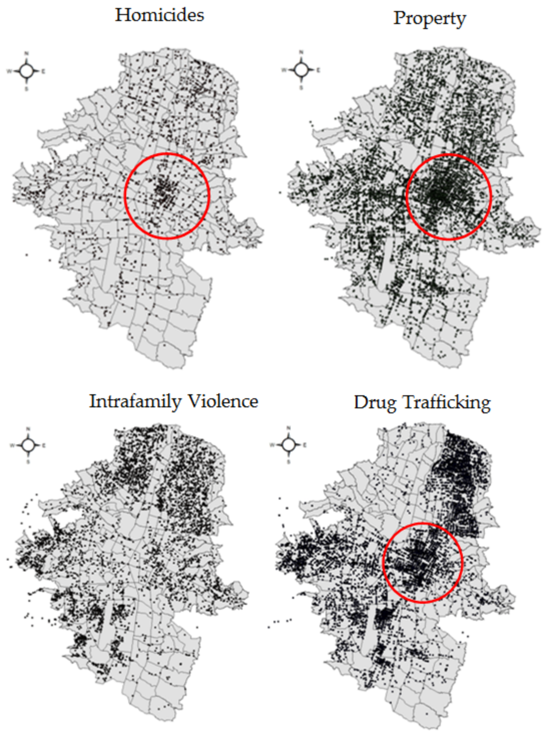


Figure A2: Distribution of Age at Arrest (Males)



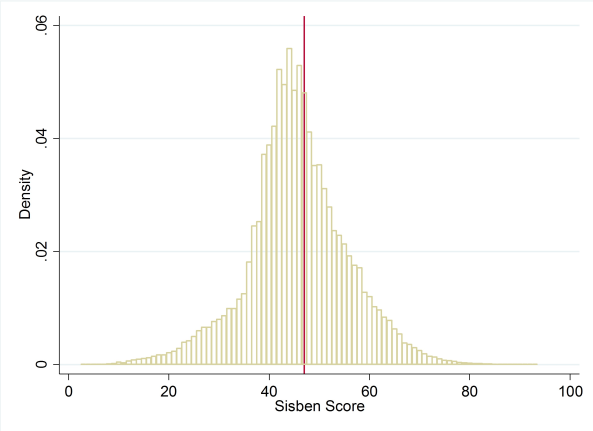
Source: Policía Nacional de Colombia. Vertical lines represent ages 13 and 26.

Figure A3: Location of 'in-the-act' arrests by type of crime, 2005-2013.

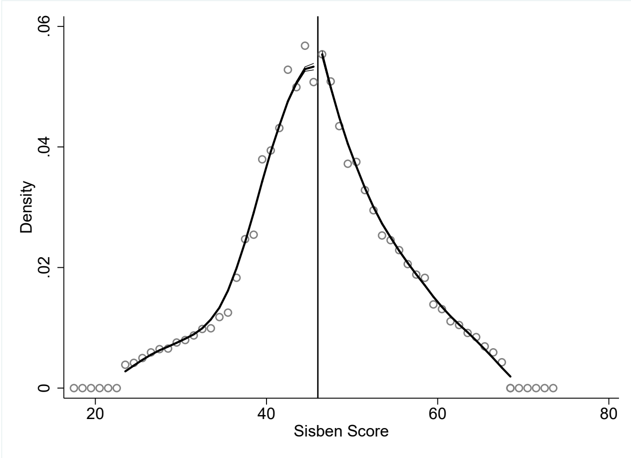


Source: Medina and Tamayo (2011) using Policía Nacional de Colombia. Dots indicate arrests. Bold lines are neighborhood boundaries. The red circle specifies the downtown of the city.

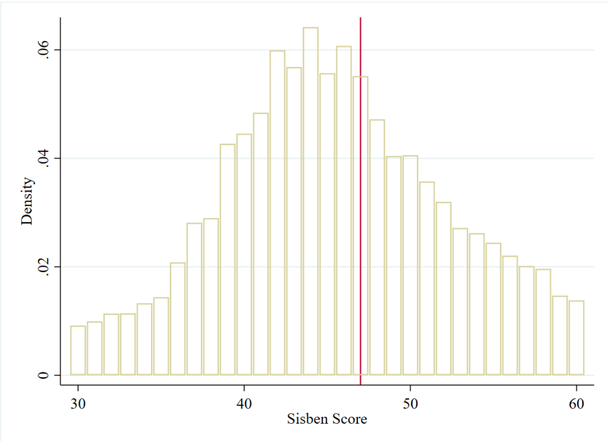
Figure A4: Sisben score Distribution



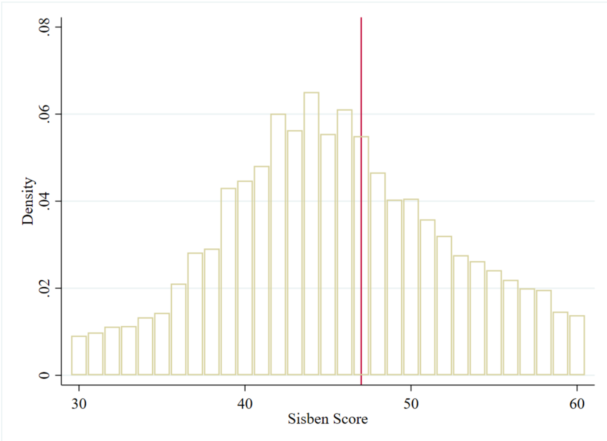
(a) Score distribution for males



(b) McCrary (2008) test



(c) Distribution for males (short bandwidth)



(d) Distribution for full sample (short bandwidth)

Source: *Sisben* survey of 2002. Figures A4a-A4c includes all males (i.e. both non-criminals and arrested individuals). Figure A4d is for the full sample (not just males). Figure A4b conducts a McCrary (2008) test.

Table A1: Top 25 Crimes (of 103) by Data-driven LACE Classifications

Crime	Type	Gang Flags	Neighborhood Method
Drug Consumption / Possession	Drug	No	No
Drug trafficking / Distribution - Marijuana	Drug	No	No
Drug trafficking / Distribution	Drug	Yes	Yes
Drug trafficking / Distribution - Cocaine paste	Drug	Yes	Yes
Drug trafficking / Distribution Heroin	Drug	Yes	Yes
Use of Fake Identification, false document	Property	No	No
Motor vehicle theft (Motorcycles)	Property	No	No
Receiving Bribes (as officials)	Property	No	No
Copyright/Fraud	Property	No	No
Identity Theft	Property	No	No
Fraud	Property	No	No
Theft / Assault	Property	Yes	Yes
Robbery (To Businesses, firms)	Property	Yes	Yes
Property Vandalism	Property	Yes	Yes
Motor Vehicle Theft - Cars	Property	Yes	Yes
Burglary	Property	Yes	Yes
Simple Assault/Battery	Violent	No	No
Rape/Sexual Assault	Violent	No	No
Conspiracy to commit murder	Violent	Yes	No
Homicide	Violent	Yes	Yes
Extortion	Violent	Yes	Yes
Assault / Battery - Against Police	Violent	Yes	Yes
Manufacture, Trafficking Firearms / Weapons	Violent	Yes	Yes
Intimidation and Stalking	Violent	Yes	Yes
Terrorism	Violent	Yes	Yes
Kidnapping	Violent	Yes	Yes

List of top crimes by type and enterprise classification, out of 103 crimes. LACE crimes are those "likely associated with criminal enterprises," and as such most reflective of individual sorting into criminal occupations. Non-LACE crimes are the remaining, more likely representing crimes of impulse or opportunity. The 'Gang Flags' method lists whether the crime has a high propensity to receive a police reported flag of known gang affiliation at the time of arrest. The 'Neighborhood Method' classifies crimes that have a high propensity to be in neighborhoods known to have high gang activity.

Table A2: Formal Employment By Gender

Bandwidths:	4	6	10
Panel A: Men Formal Employment in 2009			
Enrolled in SR	-0.0539***	-0.0411***	-0.0301***
	(0.0166)	(0.0103)	(0.00811)
Number of observations	133,067	180,742	247,886
Panel B: Women Formal Employment in 2009			
Enrolled in SR	0.00560	-0.0130*	-0.0169*
	(0.00757)	(0.00786)	(0.00889)
Number of observations	156,942	213,755	292,980

Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. We use the *Sisben* survey of 2009 to construct formal employment. Tables report two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff. Regressions control linearly for the 2002 Sisben score, flexibly around the cutoff. We cluster standard errors by *comuna*.

Table A3: Baseline (2002 *Sisben* Survey) balance tests

	Whole Sample	Male Youth
First Principal Component	0.110	-0.00246
	(0.0917)	(0.0446)
Years of Education	0.00248	0.0275
	(0.0606)	(0.0552)
Age	0.0255	0.00976
	(0.0300)	(0.0113)
Age Specific Education Gap	0.0159	-0.0208
	(0.0602)	(0.0600)
HH Head Years of Education	-0.0832	-0.0560
	(0.0592)	(0.0583)
Unemployed	0.0113	0.0111
	(0.00738)	(0.00697)
Married	0.0275	0.0254
	(0.0175)	(0.0175)
Employed	-0.0115	-0.0139
	(0.0102)	(0.0106)
Attending School	0.000756	0.000788
	(0.00270)	(0.00271)
Socieconomic Stratum 2	-0.0129	-0.00477
	(0.0144)	(0.0112)
Socieconomic Stratum 1	0.0236*	0.00362
	(0.0126)	(0.00916)
Own House	0.0218	0.00752
	(0.0127)	(0.0110)
Less than 6 years Olds	0.0131	0.0122
	(0.0129)	(0.0126)
HH Head Age	0.000756	0.338**
	(0.00270)	(0.136)
Observations	246,974	28,675

Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. Tables report reduced form coefficients on being below the Sisben cutoff, where Sisben score is measured in 2002. Regressions are for the optimal bandwidth of 6 percentage points. Regressions control linearly for the Sisben score, flexibly around the cutoff. All variables are measured in 2002. We cluster standard errors by *comuna*. Neighborhood strata indicate the official socioeconomic strata of the neighborhood. First Principal Component takes the first principal component of all other variables.

Table A4: Non-LACE Crimes: Heterogeneity by Comuna

Bandwidths:	4	6	10
Panel A: Non-LACE Violent Crimes			
Enrolled in SR	0.00886	0.00333	-0.00421
	(0.0167)	(0.0114)	(0.0123)
Enrolled* Gang Comuna	0.00773	0.00236	0.0126
	(0.0240)	(0.0150)	(0.0144)
F stat	92.6	134.6	204.3
Number of observations	18,419	24,768	33,702
Panel B: Non-LACE Property Crimes			
Enrolled in SR	-0.00571	-0.00526	-0.00503
	(0.0168)	(0.0120)	(0.00969)
Enrolled* Gang Comuna	-0.0176	-0.0122	-0.0100
	(0.0180)	(0.0153)	(0.0122)
F stat	100.2	161.8	241.4
Number of observations	18,240	24,523	33,358
Panel C: Non-LACE Drug Crimes			
Enrolled in SR	-0.0379	-0.0494	-0.0292
	(0.0299)	(0.0307)	(0.0222)
Enrolled* Gang Comuna	0.0132	0.0155	0.00281
	(0.0292)	(0.0211)	(0.0195)
F stat	94.8	135.1	197.3
Number of observations	19,150	25,740	35,104

Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. Non-LACE, as determined by the data-driven classifications summarized in Table A1, are those more likely representing crimes of impulse or opportunity rather than activity of criminal enterprises. Tables report two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff and an interaction between high-gang comunas and being below the cutoff. The Sisben score is measure in 2002, SR enrollment in 2005, and crime outcomes are measured between 2005 and 2013. Regressions include comuna fixed effects and an interaction between high-gang comunas and indicators for SR enrollment. Regressions control linearly for the Sisben score, flexibly around the cutoff. We consider only males between 21 to 26 years old in 2013. We cluster errors by comuna. The mean arrest rate across all five gang comunas are 18%, which is also the mean arrest rate in low-gang comunas.

Table A5: Neighborhood Classification Method

Bandwidths:	4	6	10
Panel A: LACE Violent Crimes			
Enrolled in SR	0.0171**	0.0121*	0.00891**
	(0.00751)	(0.00684)	(0.00387)
Number of observations	17,995	24,198	32,931
Panel B: LACE Property Crimes			
Enrolled in SR	0.0335**	0.0271**	0.0192*
	(0.0131)	(0.0122)	(0.0107)
Number of observations	18,426	24,740	33,625
Panel C: LACE Drug Crimes			
Enrolled in SR	0.0284*	0.0108	-0.00197
	(0.0163)	(0.0126)	(0.0115)
Number of observations	18,909	25,447	34,661

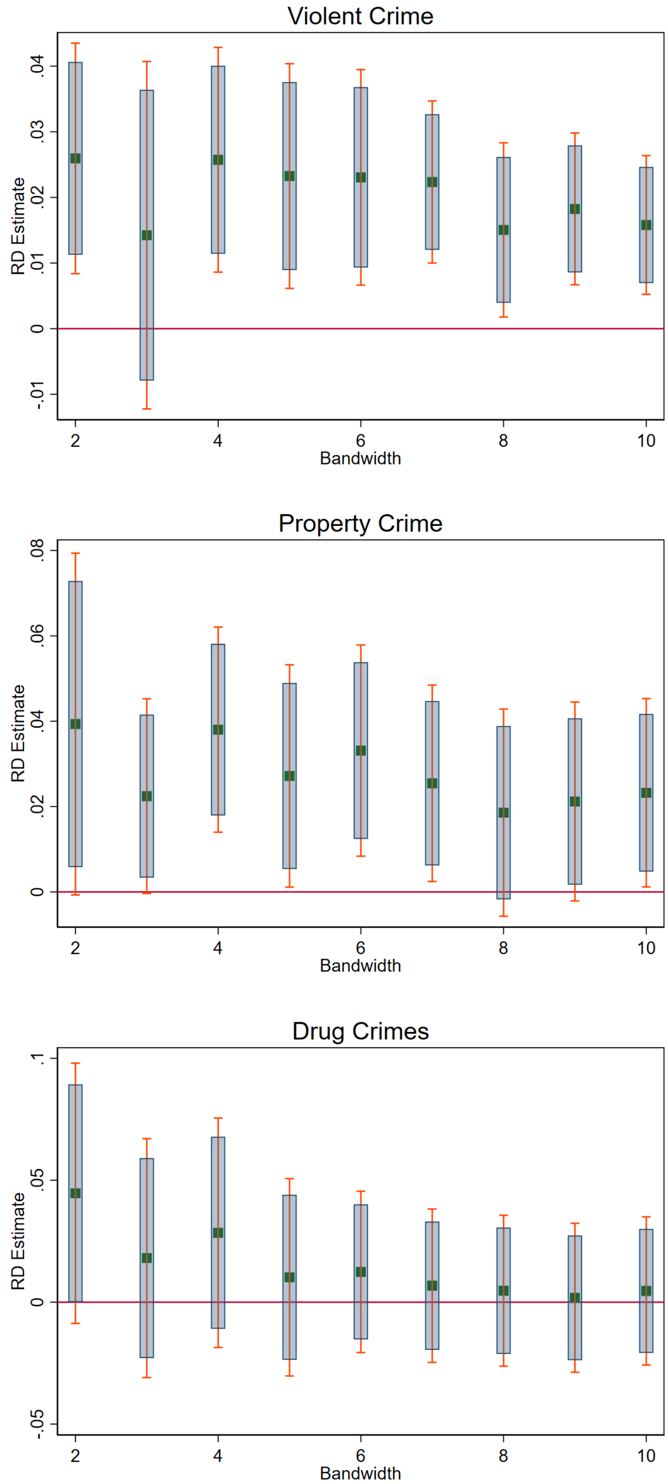
Note: Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Tables report two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff. The Sisben score is measured in 2002 and SR enrollment in 2005. Crime data are from 2005 to 2013. Results use the neighborhood classification method described in the text to classify crimes. Regressions control linearly for the Sisben score, flexibly around the cutoff. We consider only males between 21 to 26 years old in 2013.

Table A6: Semi-parametric RD with Bias Correction

Type of Crime	Violent	Property	Drug
Enrolled in SR	0.0164	0.02794	0.00768
Standard error	(0.00972)	(0.01636)	(0.02069)
Bias corrected p-value	0.077	0.052	0.57
Bandwidth	5.2	5.8	6.6
Bias correction bandwidth	9.9	8.9	9.6
Number of observations	24,206	26,511	29,102

Note: Results using the Calonico et al. (2014a) CCT method for estimation, where the primary estimation uses a linear functional form and the bias correction uses a quadratic form. Arrests are restricted to LACE crimes, which are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Tables report fuzzy RD two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff, where Sisben score is measured in 2002. Crime data is from 2005 to 2013. We consider only males between 21 to 26 years old in 2013.

Figure A5: Robustness to Bandwidths (LACE Crime)



Note: Coefficients of RD 2SLS regressions where the first stage is SR Enrollment on being below the Sisben cutoff. Sample of LACE crimes only. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Grey bars indicate 90% confidence intervals. Red lines indicate 95% confidence intervals.

Table A7: Robustness Check: Including Other Crimes in the Sample

Bandwidths:	4	6	10
Panel A: LACE Violent Crimes			
Enrolled in SR	0.0213***	0.0195***	0.0136***
	(0.00718)	(0.00676)	(0.00416)
Number of observations	21,720	29,235	39,877
Panel B: LACE Property Crimes			
Enrolled in SR	0.0320***	0.0289***	0.0204**
	(0.0100)	(0.00998)	(0.00894)
Number of observations	21,720	29,235	39,877
Panel C: LACE Drug Crimes			
Enrolled in SR	0.0238	0.0107	0.00405
	(0.0195)	(0.0138)	(0.0129)
Number of observations	21,720	29,235	39,877

Note: The sample includes other crimes. For instance, when looking at violent enterprise-crime arrests as the outcome of interest, property crime, drug crime and violent non-enterprise crime arrests are also in the sample grouped with the people never arrested in this period. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. Tables report two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff. Regressions control linearly for the Sisben score, flexibly around the cutoff. We consider only males between 21 to 26 years old in 2013.

Table A8: Robustness Check: Including Repeat Arrests

Bandwidths:	4	6	10
Panel A: LACE Property Crimes			
Enrolled in SR	0.0449***	0.0414***	0.0284**
	(0.0123)	(0.0145)	(0.0128)
Number of Observations	18,488	24,822	33,739
Panel B: non-LACE Property Crimes			
Enrolled in SR	-0.0158	-0.0136	-0.0118
	(0.0217)	(0.0172)	(0.0115)
Number of Observations	18,265	24,558	33,409
Panel C: LACE Violent Crimes			
Enrolled in SR	0.0175**	0.0154*	0.0101*
	(0.0088)	(0.0088)	(0.0052)
Number of Observations	18,071	24,294	33,060
Panel D: non-LACE Violent Crimes			
Enrolled in SR	0.0145	0.0066	0.0016
	(0.0176)	(0.0123)	(0.0144)
Number of Observations	18,480	24,842	33,800
Panel E: LACE Drug Crimes			
Enrolled in SR	0.0369	0.0090	-0.0028
	(0.0270)	(0.0191)	(0.0177)
Number of Observations	18,524	24,944	33,973
Panel F: non-LACE Drug Crimes			
Enrolled in SR	-0.0336	-0.0585	-0.0428
	(0.0310)	(0.0322)	(0.0245)
Number of Observations	19,327	25,982	35,450

Note: The sample includes repeat arrests. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Non-LACE crimes are the remaining, more likely representing crimes of impulse or opportunity. Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. Tables report reduced form coefficients. Regressions control linearly for the Sisben score, flexibly around the cutoff. We consider only males between 21 to 26 years old in 2013. We do not include controls.

Table A9: Simultaneously both Informal (in 2009) and Arrested

Bandwidths:	4	6	10
Panel A: LACE Violent Crimes			
Enrolled in SR	0.0104	0.00997**	0.00990**
	(0.00771)	(0.00405)	(0.00436)
Number of observations	12,015	16,023	21,733
Panel B: LACE Property Crimes			
Enrolled in SR	0.0393**	0.0331***	0.0183
	(0.0184)	(0.0122)	(0.0128)
Number of observations	12,244	16,319	22,074
Panel C: LACE Drug Crimes			
Enrolled in SR	0.0207	0.0299***	0.0291**
	(0.0177)	(0.00983)	(0.0126)
Number of observations	12,253	16,368	22,199

Note: The outcome is arrests only for those not formally employed as measured in 2009. We exclude all arrests post 2009. Standard errors in parentheses. *** significant at 1%; ** significant at 5%; * significant at 10%. LACE crimes are those "likely associated with criminal enterprises," as determined by the data-driven classifications summarized in Table A1, and as such most reflective of individual sorting into criminal occupations. Tables report two-staged least squares (2SLS) coefficients where the first stage is SR enrollment on being below the Sisben cutoff. The Sisben score is measured in 2002, and SR enrollment in 2005. Crime is measured between 2005 and 2009. Regressions control linearly for the Sisben score, flexibly around the cutoff. We consider only males between 21 to 26 years old in 2013.