Criminal Records, Family Formation and Violence Against Women *

Jeffrey Grogger† Hans Grönqvist‡ Susan Niknami§

March 26, 2026

Abstract

Policies that expand public access to criminal records are often justified as a way to protect potential victims. But making such information widely available may also change who forms relationships, which relationships survive, and how exposed men behave. This paper studies the causal effects of public access to criminal records on partnership formation and violence against women. We exploit the unexpected launch of an online criminal background check platform in Sweden in 2014, which fundamentally changed criminal background checks by allowing the general public to search, anonymously and free of charge, for all individuals charged with any crime. The platform attracted widespread attention and generated roughly 100 million searches by 2019. To isolate the informational content of criminal records, we exploit court-specific disclosure cutoffs created by administrative rules that prevented the platform from accessing records older than roughly five years. Using population-wide administrative registers with complete criminal histories, we compare individuals whose records became publicly visible with otherwise similar individuals whose charges fell just outside the disclosure window.

Our results suggest that the same disclosure policy can reduce women's exposure to risky partners along one margin while increasing violence along another. We find substantial heterogeneity by baseline partnership status. Among men partnered at baseline, public listing reduces the probability of remaining partnered by 7.8 percentage points, with especially large effects in newly formed relationships and in couples where the woman has greater baseline bargaining power. We find no statistically significant effect on violence against women in this group, but we do find improvements in women's labor market outcomes, consistent with disclosure enabling some women to leave lower-quality or risky relationships. By contrast, among men unpartnered at baseline, public listing reduces the

*We are grateful to comments from seminar participants at Erasmus University (Rotterdam) and Linnaeus University. Part of this work was completed while the authors visited the CEP at LSE. This work is funded by the Swedish Research Council and Handelsbankens forskningsstiftelser.

†University of Chicago, London School of Economics, NBER, and IZA. jgrogger@uchicago.edu

‡IFN, Linnaeus University, CESifo, RFBerlin. hans.gronqvist@ifn.se

§Stockholm University. susan.niknami@sofi.su.se

probability of entering a partnership by 8.7 percentage points and increases the probability of committing violence against women by 4.1 percentage points. These effects do not appear to be driven by changes in reporting behavior or by labor market deterioration. Instead, public listing increases prescriptions associated with stress-related conditions and alcohol or drug misuse, and it shifts matching toward women with a higher predicted risk of future domestic abuse. The increase in violence is also muted among men treated for depression before listing, pointing to an important role for psychological vulnerability and stress. The results reveal a core trade-off in public criminal records disclosure: access improves screening and helps some people exit relationships, but it also deepens exclusion and distress among unpartnered men, with downstream effects on violent behavior. More broadly, the findings highlight the consequences of making sensitive personal information searchable at scale.

KEYWORDS: *Background checks*; *Divorce*; *IPV*; *Family spillovers*; *Stigma*;
JEL CODES: K42; J12; J16

1 Introduction

How does newly searchable information about past criminal offending reshape the market for intimate partners? Policies that expand public access to criminal records are often justified as a way to protect potential victims. But making such information widely available may also change who forms relationships, which relationships survive, and how exposed men behave. In standard models of marriage and household formation, information about partner quality affects both match formation and relationship stability (Becker, 1973, 1974; Chiappori et al., 2017). Criminal histories are especially salient because disclosure may both help women avoid risky partners and induce behavioral responses among exposed men, with consequences for violence against women (Aizer, 2010; Card and Dahl, 2011; Anderberg et al., 2016).[1] This paper studies the consequences of public criminal-record disclosure for partnership formation, partnership dissolution, and violence against women by exploiting the unexpected launch of a website that made criminal court records free, anonymous, and searchable online.

The context is Sweden where, in January 2014, the private firm Lexbase launched a website that allowed users to search whether a specific person appeared in criminal court records. Before 2014, those records were formally public but costly to obtain in practice: retrieving them typically required contacting local courthouses, so private background checks were cumbersome and rare. Lexbase changed this overnight by making criminal-history information cheap, immediate, and anonymous to access. Unlike the few criminal background-check platforms in other countries, which typically restrict access to particular users or offenses, Lexbase provided broad access to all criminal charges brought in court. The launch was heavily publicized and rapidly adopted, generating 32 million searches in its first year.

We identify the causal effect of public disclosure by exploiting administrative rules that sharply limited Lexbase's access to older cases. Because records older than about five years were effectively unavailable to the platform, defendants whose cases fell just within the disclosure window were substantially more likely to appear in the database than otherwise similar defendants whose cases fell just outside it. Using population-wide Swedish administrative data with complete criminal histories and rich information on partnerships, labor-market outcomes, prescriptions, and violence, we implement a fuzzy regression discontinuity design around court-specific cutoffs. This design isolates the informational effect of public listing from the direct effects of criminal charges and formal sanctions.

[1] In 2017, nearly 40 percent of dating couples indicated that they had met online (Rosenfeld et al., 2019). Prior to dating, roughly half conducted a search of their online match, either via social media or a search engine (Avast, 2021). Nearly one in 12 reported running a criminal background check, and one-third of adults who had been in a romantic relationship admitted to stalking their partner online without consent (NortonLifeLock and Poll, 2022).

We first verify that baseline characteristics of defendants, including original crime type and sentence severity, are balanced for individuals charged just before and just after the cutoff. We then show that being sentenced after the cutoff increases the probability of being publicly listed in the database by 0.719 (s.e. = 0.029).

Our results show large effects on partnership outcomes, with substantial heterogeneity by baseline relationship status. Among men who had a partner when Lexbase launched, public listing increases the probability of separation by 7.8 percentage points. The effect is especially pronounced in recently formed relationships and in couples where women had greater baseline bargaining power, as proxied by relative earnings or schooling. Separations occur quickly, with most of the effect concentrated in the first two years after launch. For this group, however, we find no statistically significant increase in violence against women. We do find improvements in women's labor-market outcomes, consistent with disclosure enabling some women to leave relationships they value less or perceive as risky.

The effects differ sharply for men who were unpartnered at baseline. For these men, public listing reduces the probability of forming a partnership by 8.7 percentage points and increases the probability of committing violence against women by 4.1 percentage points. The increase in violence is no larger for charges resulting in conviction than for charges that were dismissed, making it unlikely that the result is driven solely by changes in reporting or by underlying differences in culpability. Nor do we find meaningful effects on men's employment or earnings. Instead, public listing increases prescriptions associated with stress-related conditions and with alcohol- and drug-related problems. It also shifts matching toward more vulnerable women: conditional on subsequently partnering, listed men become more likely to match with women who, based on predetermined characteristics, have a high predicted risk of future domestic abuse. These patterns point to stigma, psychological distress, and re-sorting in the partner market as central mechanisms.

The results therefore point to an important asymmetry. Public disclosure appears to improve screening and facilitate exit from some existing relationships, but it also increases exclusion and distress among men who are not yet matched, responses that can in turn raise the risk of subsequent violence.

This paper contributes to four strands of literature. First, it contributes to research on the economics of the family by showing that an information shock about partner quality—separate from income shocks, incarceration, or changes in bargaining institutions—has large causal effects on union formation, union dissolution, and violence against women (Becker et al., 1977; Aizer, 2010; Chiappori et al., 2017; Anderberg et al., 2016). Relative to existing work, our setting allows us to hold constant much of the underlying criminal history and formal sanction while varying whether that history becomes publicly visible.

Second, the paper contributes to the literature on the collateral consequences of criminal records. Existing economics research has focused mainly on employment, hiring, and the effects of policies that restrict employer access to criminal history information (Holzer et al., 2006a; Agan and Starr, 2018; Doleac and Hansen, 2018; Craigie, 2019; Jackson and Zhao, 2016; Rose, 2021). We show that the consequences of criminal-record disclosure extend well beyond the labor market: they reshape household formation, partner sorting, and women's safety. More broadly, our estimates isolate the informational content of a criminal record, complementing work that studies the effects of incarceration or conviction themselves (Lopoo and Western, 2005; Apel, 2013; Fallesen and Andersen, 2017; Siennick et al., 2014; Charles and Luoh, 2010; Svarer, 2011).

Third, the paper contributes to the economics of violence against women. Existing work has emphasized changes in household resources, bargaining power, and institutional responses (Aizer, 2010; Anderberg et al., 2016; Bhuller et al., 2024; Calabresi and Rodríguez-Planas, 2025). We identify a previously unexplored margin: the public disclosure of criminal histories. Our results suggest that the same disclosure policy can reduce women's exposure to risky partners along one margin while increasing violence along another.

Fourth, the paper contributes to a broader literature on information disclosure and searchable personal information. Digital technologies increasingly transform information that is legally public but practically obscure into information that is instantly searchable, reducing information frictions while creating new privacy trade-offs (Goldfarb and Tucker, 2019; Acquisti et al., 2016). In matching markets, greater transparency can improve screening and match quality, but may also generate exclusion and other inefficiencies (Akerlof, 1970; Ostrovsky and Schwarz, 2010). Our setting provides causal evidence on this trade-off in a high-stakes environment. We show that making criminal histories widely searchable improves screening and facilitates exit from some relationships, but also increases exclusion and behavioral responses among unmatched individuals, leading to unintended consequences in the form of increased violence against women. These results highlight how the welfare effects of disclosure depend on which side of the market is affected.

The rest of the paper proceeds as follows. The next section describes the Swedish institutional setting and the launch of Lexbase, and explains how record-retention rules generate our regression discontinuity design. We then describe the data, present the main results, explore mechanisms and heterogeneity, and conclude.

2 Background

2.1 Prior literature on the economics of marriage

The empirical literature on the economics of marriage began with Becker et al. (1977), who analyzed marriages in cross-sectional data. They found that divorce probabilities were higher at both higher and lower levels of income. Since variation in the tails is more likely to be the effect of surprises than it is in the middle of the distribution, they inferred that their results were due to shocks, but could not rule out a role for match quality. Weiss and Willis (1997) analyzed divorce rates using longitudinal data, potentially letting them control for initial match quality, but focused on differences between income realizations and predictions from an auxiliary earnings regression, rather than true shocks. Charles and Stephens (2004) also analyzed longitudinal data, focusing on divorce responses to layoffs. This approach may distinguish the effects of shocks from the effects of match quality if the layoffs are unanticipated. Two studies, by Cesarini et al. (2023) and Hankins and Hoekstra (2011) study lottery winners. Such winnings have the advantage of providing exogenous variation in income shocks, but they are relatively rare and involve only individuals who participate in lotteries.

Although much of the empirical work on the economics of marriage has focused on interactions between marriage and the labor market, other factors may affect partners' perceptions of their gains from union as well. Charles and Stephens (2004) observe that layoffs increase divorce risk, but that disabilities with similar implications for income do not. This leads them to conclude that surprises about non-economic aspects of one's partner's suitability as a mate may be more important than the financial losses entailed by a job loss in raising divorce risk.

2.2 Prior literature on crime and family formation

One such non-economic aspect of a partner's suitability may entail their law-abidingness, or lack thereof. A handful of studies focuses on links between crime, dating, and marriage.[2] Svarer (2011) analyzes population register data from Denmark, and reports that criminal convictions reduce the likelihood of forming a union with a partner from a better-off family. van Schellen et al. (2012) study official records from the Netherlands, and find that convictions both reduce the likelihood of marrying, and among those who marry, raise the likelihood of marrying someone else with a criminal record. They also

[2]We distinguish the household-level studies discussed here from several market-level studies that estimate how widespread incarceration, by reducing the sex ratio, affects marriage. These studies include Charles and Luoh (2010), Liu (2019) and Caucutt et al. (2021).

report that dissolution risk is higher among men with convictions. Evans et al. (2025) conducted a street-intercept survey, asking respondents whether they would be willing to date prospective matches with various types of criminal records. They report that most respondents were seriously opposed to dating someone with a criminal record, especially for a serious offense. Evans and Vega (2021) conducted a correspondence experiment on a online dating service to gauge how prospective matches reacted to information about a criminal record on one's profile. Results were curiously positive for white males, but negative for Black males.

Most of the findings from both of these literatures are broadly consistent with what Becker's model, in some cases augmented for a criminal past, would predict. However, most struggle to disentangle assortative mating on observable characteristics, or the role of shocks in dissolutions, from unobserved aspects of match quality. The studies that are most successful along these lines tend to have been drawn from special populations, to be based on small or potentially unpresprentative samples, or to analyze stated preferences rather than actual behavior. In our analysis below, we make use of data from national population registers, observe actual formations and dissolutions of unions, and deploy an RDD to deal with potential confounding due to observed match quality.

2.3 Criminal background checks

In most countries, criminal background checks are used by employers, landlords, educational institutions and the like to address concerns surrounding public safety. These checks may mitigate potential liability, particularly in the realms of negligent hiring or negligent renting doctrines.[3] Background checks were historically provided by public agencies, but increasingly are provided by private companies, especially in the United States. A background check typically reports the individual's name associated with the record, the jurisdiction of origin, the record's creation date, and a case number or law enforcement identifier (NLCS, 2019). While a brief description of the offense, such as "possession of marijuana", may be provided, details about the offense or related circumstances are usually omitted. Some public sources only display criminal convictions, whereas private firms may report arrests, some of which later result in dismissal (Bushway and Sweeten, 2007).

[3]About 94 percent of employers in the United States conduct some form of criminal history check, and about 90 percent of landlords run background checks on prospective tenants (NLCS, 2019) In 2004, a survey from the United States showed that about 50 percent of the employers check the criminal background of job applicants (Holzer et al., 2006b)

2.3.1 The Swedish setting before Lexbase

Regulations governing criminal record checks in Sweden were traditionally rather stringent. Historically, they were primarily available only to employers (Backman, 2012). Even then, employers were limited in their ability to request a candidate's criminal record, which involved a time-consuming process that was typically permitted only at the concluding stages of the hiring process. As a result, survey data from a large sample of employers collected in 2011 showed that only approximately 14 percent of the 1,200 employers reported to have engaged in such practices (SOU, 2011).

2.3.2 The Lexbase website

Lexbase was established as a business in 2012. Between then and the January 2014 public launch of its website, the company began acquiring criminal background data through direct engagement with individual courts, requesting access to criminal charges stored in the digital administrative system known as Vera. A government decree mandated that the courts honor Lexbase's requests.[4]

With the launch of its website in January 2014, Lexbase completely transformed the landscape of criminal background checks in Sweden. Instead of targeting solely employers, the platform offered unrestricted access to criminal records to the general public. The website allowed anyone to anonymously search individuals by name, leveraging data sourced from the records maintained by the courts. The search algorithm links the name of the searched individual to their current age and place of residence and reveals whether that person has been charged with a crime. This means that even for common names it is typically possible to uniquely identify an individual.[5]

Unlike many other online criminal databases, which specialize in specific crime types or cater exclusively to certain sectors such as employers or landlords, the Lexbase platform contains a comprehensive repository of every criminal charge in Swedish courts. The basic free search service simply indicates whether the subject of the search has previously been charged with a crime. For a nominal fee, at the time about 5 USD, registered users may also obtain full court proceedings, gaining access to information about the number and type of charges, their disposition, sentence information, and the address and personal

[4]See decree (2001:639).

[5]It is also possible to search using the personal identity number (cf. SSR). The personal identity number is a unique identifier assigned to each resident at birth or upon immigration. It is used across virtually all areas of society—healthcare, taxation, banking, education, employment, social insurance, and many private services. Its use is universal and deeply integrated into both public and private systems, making it one of the most widely used personal identification numbers in the world. In Sweden, anyone can request a person's personal identity number from the Tax Agency by providing the individual's full name, unless the information is protected by confidentiality rules.

id number (cf. social security number) of the individual.[6]. Appendix Figures A.1 and A.2 present the website as it was seen by the users.

Although the information provided by the website in theory was accessible through public records even before 2014, the process was laborious, necessitating visits to multiple courts to potentially find a matching name. The introduction of this particular platform represented a significant reduction in the cost of accessing criminal background information, especially for the general public.

The platform immediately received large media attention and gained immense popularity. Data provided by the company behind the platform show that 32 million background checks were conducted in just the first year. Appendix Figure A.3 shows that by 2019 when the first competing platform was introduced, the cumulative number of searches had soared to 103 million. Appendix Figure A.4 shows the sharp increase in the number of media articles written in printed press mentioning the database. Appendix Figure **??** shows that from the week after its inception and up until 2019, Lexbase consistently outnumbered other popular Google search terms in Sweden, surpassing common queries such as "new car" or "best restaurants". Media coverage primarily focused on how the platform could be used to conduct background checks, often accompanied by anecdotal accounts from individuals concerned that the public revelation of their criminal history would adversely affect their future outcomes.[7]

Since the platform did not have any competitors until 2019, the massive use of the database cannot be explained by a shift in the use from other similar platforms. It also seems unlikely that even employers substituted away from official background checks by the police to instead using the platform. First, the introduction of the platform did not coincide with any meaningful shift in the long-run trend towards increased use of official criminal background checks (see Appendix Figure A.7). Secondly, the demand for regular background checks was relatively low among employers. As discussed earlier, a large survey of 1,200 firms just before the introduction of the platform revealed that 14 percent of employers reported that they check criminal records (SOU, 2011) and only two percent of the job ads posted on the Public Employment Office's mentioned criminal background checks. Third, the number of official background checks, 171,229 in 2011, is completely dwarfed by the about 20 million of searches conducted every year on the platform. Taken together, the large use of the platform is most likely explained by a high demand for the service in the general public.

The platform was shut down by its internet service provider just two days after its launch,

[6]Despite its low cost, acquiring the complete court proceedings seems to have been a rare occurrence. Data from the database's company indicate that 0.16 % of all background searches actually resulted in a purchased court proceeding in the first year (see Appendix Figure A.3)

[7]`https://www.expressen.se/nyheter/deras-liv-forstors-av-lexbase/`

following intense criticism from Swedish media and the general public for making criminal records too easily accessible. However, between April 2014 and February 2025, the platform remained fully operational. Several attempts by the Swedish government to shut down the platform failed, as the information it provided was protected under the Swedish Constitution. However, following a ruling by the Swedish Supreme Court stating that public court documents containing personal data—such as criminal records—should not be disclosed when the purpose is to distribute them through databases accessible to paying customers, the platform no longer offers this service.

3 Data and sample construction

3.1 Administrative data

Our analysis uses population-wide register data covering the entire Swedish population from 2000 to 2018. One register contains demographic and socioeconomic information, including age, place of birth, gender, educational attainment, employment status, and income measures. Parents are linked to children using the Multigeneration Register.

Our measure of partnership status follows the definition used by Statistics Sweden: being married or cohabiting with shared children. This measure is obtained from the LISA register. Partnership formation is thus defined as either marriage or cohabiting partners having a common child, while partnership dissolution is defined as divorce or separation among cohabiting partners. The dependent variables are indicators equal to one if an individual is partnered or unpartnered in any year between 2014 and 2018.

We link these data to the Swedish Conviction Register maintained by the National Council for Crime Prevention (*Brottsförebyggande rådet*—BRÅ). The conviction register contains detailed information on criminal charges during this period, including the date and subsequent disposition. This information originates from Swedish district courts, which are courts of first instance.

A single criminal charge may encompass multiple offenses, and we observe all offenses within a given charge.[8] Violence against women is recorded when a prosecutor suspects a male of committing violence against a female in a close relationship. This definition includes violence against both current and former partners, as well as adult female relatives. The dependent variable is an indicator equal to one if violence against women occurs in any year between 2014 and 2018.

[8]The data exclude minor offenses such as speeding tickets but include offenses such as driving without a license and DUI.

To construct our estimation sample, we limit attention to individuals charged with a crime between 2007 and 2010 whose cases fall within a 16-week window around the relevant court-specific disclosure cutoff. We restrict the sample to first-time defendants and, to ensure comprehensive tracking of outcomes, further restrict it to individuals aged 20 to 60 in the year of the initial charge.[9]

3.2 Summary statistics

Table 1 reports sample means for groups of defendants stratified by whether they appear in the database (yes/no) and whether they were sentenced after the cutoff (yes/no). These baseline characteristics are measured in the year prior to the initial charge.

Appearance in the database is endogenous, as defendants who recidivate are more likely to be included when the database is launched. Sentencing after the (court-specific) cutoff serves as the instrument. As expected, differences in sample means are smaller when comparing defendants sentenced after versus before the cutoff, and, as we show later, these differences are not statistically significant. In contrast, differences between defendants who appear in the database and those who do not are larger and also statistically significant.

4 Retention limits and the RDD

4.1 Retention limits and cutoff dates

A critical element of our RDD is the legal retention limit for criminal records, which in combination with the Lexbase launch date, determines assignment to treatment. In principle, courts are required to expunge criminal records after five years.[10] This requirement pertains not to Lexbase, but to the underlying Vera system from which Lexbase obtains data. Thus in principle, Lexbase's requests should have resulted in criminal records extending back five years from the date the request was made.

Figure 1 helps to illustrate our design. It depicts the timeline for a court whose records were requested in January 2014 when Lexbase website was launched. We assume for simplicity that five years of criminal history records were made available at that time.

[9] Appendix B describes how the variables in our analysis are constructed.

[10] The courts were still obliged to preserve cases in printed form. In order to search for a specific case a person had to visit the court to manually search for the case number in a volume of books that compiles all the cases that have been processed by the court.

With the five-year retention limit, any offenses committed prior to the cutoff date of January 2009 would not have appeared in Lexbase, and thus would not have been available to anyone visiting the website. In contrast, any offenses committed after the January 2009 cutoff date would have appeared in Lexbase upon its January 2014 release.

Our treatment group consists of all individuals who were charged with an offense within a short interval after the cutoff date. Our comparison group consists of individuals who charged within a similarly short interval prior to the cutoff date. The figure illustrates treatment and comparison groups based on an interval of 32 weeks, on which we base our main results.[11] If the intervals are short enough, the only systematic difference between the treatment and control groups should be the fact that the treatment group's charges were searchable in Lexbase, whereas the control group's were not. The RDD exploits this difference to identify the effects of a public disclosure of a criminal record on our vrious outcomes over the January 2014-December 2018 follow-up period.

In practice, Lexbase requested records from different courts at different times, so cutoff dates are court-specific. Some vagaries surrounding the retention limits, as well as the massive scale of the Lexbase requests, further contributed to variation among courts in effective cutoff dates. The specific cutoff date for each court was determined by the timing of Lexbase's request for information, the length of time it took the court to process the request, and the exact expungement procedures followed by the court.

To estimate court-specific cutoff dates, we obtained data from Lexbase on the number of cases added to the database for each court in each month. We used these data, together with nationwide register data on the near-universe of all criminal charges in Sweden since 1986, to estimate the share of cases included in the Lexbase data for each court and week.[12] We then computed the change in the share of included cases each week by calculating the difference between the average share of cases in the preceding 16 weeks and the average share in the subsequent 16 weeks. We estimated the cutoff date as the week exhibiting the largest increase in the share of included cases over the entire period.[13]

[11] We use this interval for our main results, but experiment with other intervals as well to check robustness.

[12] The criminal justice registers do not include offenses by persons without national ID numbers, who are mostly immigrants.

[13] In the empirical analysis we focus on the 32 courts with observed cutoff dates before 2010 to ensure the representation of all courts in our sample within the database at its launch. It is unclear whether all courts were actually included in the database at the time it was launched. Thus, it is possible that some courts were added at a later point in time. Additionally, we excluded four courts with no clear evidence of cutoff. These courts, which represent 7.5 percent of the total number of cases, are DC Umeå, DC Gällivare, DC Haparanda, DC Värmland.

4.2 The RDD

The discussion above motivates our RDD design, which we estimate from the sample of individuals aged 20 to 60 who were charged with a crime for the first time between 2007 and 2010 and whose cases fall within a 16-week window around the relevant court-specific disclosure cutoff. The relevant regressions can be written as

$$D_{ij} = \alpha_0 + \alpha_1 Z_{ij} + \alpha_2 W_{ij} + \alpha_3 W_{ij} Z_{ij} + \varepsilon_{ij} \tag{1}$$

$$Y_{ij} = \beta_0 + \beta_1 Z_{ij} + \beta_2 W_{ij} + \beta_3 W_{ij} Z_{ij} + u_{ij} \tag{2}$$

The variable D_{ij} is a binary indicator equal to one if individual i has a criminal record that appears in Lexbase, and equals zero otherwise. The variable Y_{ij} is an outcome, such as marital status, for the ith individual covered by the jth court. The running variable Z_{ij} is the number of days between the date individual i's was charged and the cutoff date for court j. The dummy variable $W_{ij} = 1(Z_{ij} > 0)$ equals one if i is in the treatment group and equals zero if he is in the comparison group. The terms ε_{ij} and u_{ij} are unobserved disturbance term.

The α and β terms are coefficients to be estimated. Equation (1) is the first stage regression. The coefficient α_3 shows the change in the likelihood of public disclosure just after the release date. We expect $\alpha_3 < 1$ since members of the comparison group may reoffend at some point between the cutoff date and the Lexbase launch date. In this case, they will have an offense that is visible in Lexbase on the release date and $D_{ij} = 1$.[14]

Equation (2) is the reduced form. The coefficient β_3 shows the effect of being charged (for the first time) during the period immediately following the cutoff date. Provided that W_{ij} is uncorrelated with ε_{ij}, β_3 should identify the intent-to-treat effect of a criminal charge on the outcome variable.

In practice, our regressions also control for the court-specific cutoff date. Some specifications add higher-order terms in W_{ij} and Z_{ij}, whereas others incorporate various control variables constructed from Swedish population registers.

Using this variation in the cutoff as an instrument for being included in the database, we identify the LATE (Local Average Treatment Effect) of being listed defendants are at the margin of being listed compared to defendants who were charged but not listed.

[14] We may also observe $\alpha_3 < 1$ due to possible measurement error in the estimated cutoff date, as discussed above.

Because defendants whose original case falls just outside the disclosure window can nevertheless appear on Lexbase if they commit a subsequent offense before the website launches, our two-stage least squares estimates should not be interpreted as the effect of disclosure for all offenders near the cutoff. Under monotonicity, the IV estimand instead recovers a LATE for defendants whose appearance on Lexbase is determined by whether the original case falls within the disclosure window. In substantive terms, these compliers are defendants whose earlier case would otherwise have remained practically obscure by 2014 and who did not accumulate a newer offense that would have made them visible in any event. Relative to offenders who would have appeared on Lexbase regardless, compliers are therefore likely to have lower recidivism risk. At the same time, Lexbase plausibly generated a larger informational shock for this group, since public visibility turned on the disclosure of an otherwise fading record rather than on a subsequent offense that would itself have returned the individual to public view.

The conditions necessary to interpret the two-stage least squares estimates as the causal impact of being listed are that: (1) being sentenced after the cutoff impacts the probability of being listed, (2) that being sentenced after the cutoff only impacts the outcomes through the probability of being listed, and (3) there is a monotonic impact of being sentenced after the cutoff on the probability of being listed. We now consider whether these conditions holds in our data.

4.3 Validating the research design

First-stage and monotonicity: Estimating equation (1) we find that being sentenced after the cutoff increases the probability of appearing in Lexbase rises substantially. As seen in column (2) of Table , our estimate of α_2 is 0.703 (0.028). This indicates that the likelihood of appearing in Lexbase is 70 percentage points higher for the treatment group than the comparison group. This is a very strong first stage, as indicated by the Sanderson-Windmeijer F-statistic of 654.9 and the Cragg-Donald F-statistic of 4569.7

Figure 2 displays the average share of cases included in the database around the estimated cutoff for the courts in our sample. The figure provides a visual illustration of the first stage. As expected, the share of included cases does not jump from zero to one at the estimated cutoff month. While, there is a clear discontinuity, we observe that approximately 20 percent of defendants sentenced before the cutoff are nonetheless listed in the database, partly because some of these individuals reoffended at some point between the cutoff and the release date. Conversely, about 10 percent of defendants sentenced after the cutoff are not listed. This reflects the fact that, in some cases, courts stopped reporting slightly before the formal retention period. In addition, some courts provided

information covering periods extending further back in time than the retention period.[15]

The monotonicity assumption requires that crossing the sentencing cutoff does not reduce the probability that a defendant is listed in the database. In our setting, this assumption is plausible if sentencing after the cutoff increases the likelihood that a case is listed, and does not cause any cases that would otherwise have been listed to be excluded. While monotonicity is not directly testable, the large positive first-stage discontinuity—from 0.2 to 0.9—suggests that the cutoff operates primarily by increasing listing probability, making substantial violations of monotonicity unlikely. In line with the monotonicity assumption, we find in Table that the effect of our instrument on incarceration is consistently positive and sizable in all subsamples.

Independence and exclusion: A second condition required to interpret our two-stage least squares estimates as the LATE of public disclosure is the exclusion restriction: being sentenced after the cutoff must affect outcomes only through the probability of appearing in the database. Table A.2 provides evidence consistent with this condition. Column (1) reports estimates from a regression of treatment status, D_{ij}, on the covariates listed in Table 1. Several of these variables significantly predict appearance in Lexbase, including sex, age, household structure in 2013, education, earnings, and conviction type. Jointly, the covariates are highly predictive: the p-value on the F-test is effectively zero.

Column (2) of Table A.2 assesses the exogeneity of the instrument by regressing the assignment variable, W_{ij}, on the same set of covariates. The joint F-statistic yields a p-value of 0.234. Appendix Figure A.8 shows the results from a similar exercise where the dependent variable is the predicted value of the endogenous variable "Appearing in the database" from a regression where this variable is regressed on baseline covariates. The prediction is plotted against weeks until the court-specific cutoff. As is clear, there is no significant jump around the cutoff. These results are consistent with quasi-random assignment around the cutoff.

Of course, the exclusion restriction could still fail if sentencing after the cutoff affected later outcomes through channels other than database inclusion. This assumption is fundamentally untestable, and our estimates should be interpreted with that caveat in mind. Even so, we believe the assumption is plausible in our setting. Most importantly, because neither the future introduction of the database nor its precise launch date was known at the time of the original court proceedings, there is limited scope for sentencing date relative to the cutoff to influence later outcomes except through subsequent public disclosure. Consistent with this interpretation, Appendix Figure A.9 shows no visible evidence

[15]The slight downward slope arises because individuals who offended closer to the cutoff date had a shorter period prior to the release date during which they were at risk of re-offending.

of manipulation around the threshold.[16]

5 Results

This section presents the results from our empirical analysis. We start by providing the main results and probe robustness. We then show results for relevant subgroups of the population. We end by examining some of the underlying mechanisms.

5.1 Main results

Table 2 shows estimates of the effect of public disclosure on the probability of being partnered (i.e., married or cohabiting with shared children) in any year between 2014 and 2018. The sample consists of first-time male defendants aged 20 to 60 who were sentenced within a 16-week window around the court-specific cutoff. All regressions control for year and court fixed effects, baseline covariates, and weeks-to-cutoff, allowing for varying slopes above and below the cutoff.

Panel A presents results for the full sample. The OLS estimate in column (2) shows that public disclosure significantly reduces the probability of being partnered in the subsequent five years by 6.9 percentage points. Column (3) improves identification by presenting 2SLS estimates, where being sentenced after the court-specific cutoff is used as an instrument for inclusion in the database. The results indicate a significant 9.8 percentage point reduction in the probability of being partnered due to public disclosure of criminal records. Column (4) shows that adding baseline controls to the 2SLS specification has little effect on the results.

Panel B shows results for the subsample of males who were partnered at baseline in 2013, while Panel C shows results for males who were unpartnered at baseline. The results reveal heterogeneity by baseline partnership status. Among men who were partnered at baseline, public listing leads to a 7.8 percentage point decline in the probability of remaining partnered. In contrast, among men who were unpartnered at baseline, public listing reduces the probability of entering a partnership by 8.7 percentage points.

The results in Table 2 show that the OLS estimates are systematically smaller (in absolute value) than the IV estimates. This is consistent with two non-mutually exclusive mechanisms. First, defendants whose initial case falls just outside the disclosure window may still appear on Lexbase if they commit a subsequent offense before the platform's

[16]The McCrary test is not statistically significant at the conventional levels.

launch. Thus, the IV estimand should not be interpreted as the effect of listing for all offenders near the cutoff. Rather, the LATE pertains to defendants whose appearance on Lexbase is determined by whether their initial case falls within the disclosure window. These compliers are therefore likely to have a lower recidivism risk than offenders who would appear in the database regardless. The IV thus identifies the effect of making an otherwise fading criminal record publicly searchable, rather than the effect of disclosure for chronic offenders whose public visibility is ensured by subsequent offending.

This interpretation may also help explain why the IV estimates exceed the corresponding OLS estimates in absolute value. For compliers, Lexbase generates a particularly large information shock, since their record would otherwise have remained largely obscure.[17] By contrast, the OLS-treated group includes offenders who would have become publicly visible anyway through subsequent criminal justice contact, for whom the marginal effect of listing may be smaller.

Second, the OLS estimates may be attenuated by imperfect measurement of database inclusion and by residual selection into actual listing status. Both mechanisms would tend to make the causal effect identified by the IV exceed the corresponding OLS estimate in absolute value.

Figure **??** investigates the dynamics of the effect of public disclosure on the probability of being partnered over the period 2014–2018 by presenting year-by-year estimates. Figure A shows estimates for individuals partnered at baseline, while Figure B shows estimates for individuals unpartnered at baseline. Figure A indicates that the effect is immediate for males who were partnered at baseline, with most of the decline in partnership occurring in the first year. In contrast, the effect is more gradual and increases over time for males who were unpartnered at baseline.

Appendix Table A.3 presents the corresponding estimates for females. We find no statistically significant effect of public disclosure of criminal records for females; however, the sample size is substantially smaller, which reduces the precision of the estimates.

Table 3 presents results for the effect of public disclosure on the probability of committing violence against women, defined as the male being recorded by the prosecutor—following investigation—as a plausible suspect of violence against a female intimate partner in any year between 2014 and 2018. The 2SLS estimate in column (3) shows that public disclosure significantly increases this probability by 3.8 percentage points. We again document heterogeneity by baseline partnership status. While the 2SLS estimate in column (4) shows a statistically insignificant 2.8 percentage point increase among men

[17]For other individuals, Lexbase may represent a smaller information shock, for instance because (i) their criminal involvement may already be more visible to partners, (ii) they may already carry other negative signals, and (iii) the public listing of an additional case may matter less at the margin.

who were partnered at baseline, we find a significant 4.1 percentage point increase for unpartnered males. As before, the OLS estimates are systematically smaller than the 2SLS estimates, likely reflecting both treatment effect heterogeneity and measurement error in database inclusion.

5.2 Further robustness checks and additional results

Appendix Table A.4 shows results from robustness checks and presents additional specifications. Column (1) controls for the running variable by binning weeks and flexibly accounting for it using year-by-month fixed effects (e.g., Kolesár and Rothe, 2018). Column (2) increases the bandwidth to 20 weeks, while Column (3) reduces the bandwidth to 8 weeks and estimates a local randomization RD without controlling for the running variable. Overall, the estimates remain stable and statistically significant across specifications.

Column (4) uses, as a placebo, the dependent variable measured in 2013, i.e., the year before the launch of Lexbase. These estimates are closer to zero and statistically insignificant. Appendix Figure A.10 shows the corresponding reduced-form plots for Column (4), confirming that there is no visual evidence of a discrete jump at the cutoff.

Our baseline sample consists of first-time defendants. In column (5), we expand the sample to also include defendants with multiple prior offenses. The point estimates are slightly smaller and statistically insignificant when repeat offenders are included. One possible interpretation is that public disclosure matters most when it reveals criminal history that would otherwise have remained obscure. Among first-time defendants, searchable disclosure may therefore generate a larger belief update in the partner market.

Columns (6) and (7) stratify the sample based on the severity of the initial crime. We define a misdemeanor as a crime for which the average incarceration rate is at most 1 percent among all offenders. We define a felony as a crime for which more than 1 percent of offenders receive prison sentences.[18] For defendants charged with less severe crimes, the point estimate for the probability of being partnered increases in absolute value (i.e., becomes more negative), while the estimate for violence against women approaches zero and becomes statistically insignificant. We observe the opposite pattern for individuals charged with more severe crimes: the estimate for the probability of being partnered becomes statistically insignificant, while the estimate for the probability of committing violence against women increases relative to the full sample.

One possible interpretation is that, for less severe offenses, disclosure produces a larger

[18]The average incarceration rate across all crimes in Sweden is about 13 percent.

decline in partnering but no detectable increase in violence against women, consistent with public listing primarily improving screening among prospective or current partners. For more severe offenses, by contrast, the partnership effect is smaller while the effect on violence is larger, suggesting that listing adds less new information for screening but may intensify stigma and exclusion among men with a higher baseline propensity for violent behavior. Taken together, these patterns suggest that the welfare consequences of public disclosure depend not only on whether the information is newly revealed, but also on how exposed individuals behaviorally respond to exclusion.

5.3 Mechanisms

Many mechanisms could potentially explain our results. In this section, we present evidence that helps identify the most plausible ones. For ease of exposition, we discuss mechanisms that are more likely to matter for family formation separately from those that are more likely to matter for violence against women. These mechanisms should not, however, be viewed as mutually exclusive. On the contrary, they may interact, and some may operate simultaneously across both sets of outcomes.

As documented above, the effects of public disclosure differ substantially depending on defendants' baseline partnership status. In interpreting this heterogeneity, it is useful to note that partnered and unpartnered defendants differ markedly in their baseline characteristics. Appendix Table A.5 reports descriptive statistics for these two groups. The table shows that unpartnered men are negatively selected on a range of observable characteristics. Relative to partnered defendants, they have lower educational attainment and substantially weaker labor market outcomes. For example, whereas 80.8 percent of partnered men were employed at baseline, the corresponding share among unpartnered men was only 61.4 percent. Unpartnered defendants also appear more criminally and psychologically vulnerable: 28.1 percent are re-convicted, compared with 14.7 percent of partnered defendants, and they are more likely to have been treated for depression, sleep or anxiety disorders, or ADHD.

These differences suggest that the heterogeneity in treatment effects by baseline partnership status may partly reflect differences in underlying "quality" or risk characteristics. In other words, unpartnered defendants may already be more disadvantaged, both economically and socially, before disclosure occurs. This matters for interpretation because public disclosure may have different consequences depending on an individual's starting position in the partner market. For men who are already unpartnered and negatively selected on observable characteristics, disclosure may further reduce relationship opportunities and exacerbate existing vulnerabilities. By contrast, partnered defendants may

be relatively more advantaged to begin with, and disclosure may therefore operate more through relationship dissolution and partner responses than through complete exclusion from the partner market.

At the same time, these baseline differences are unlikely to be the whole story. Even if part of the heterogeneity reflects pre-existing differences in defendant characteristics, the results also point to meaningful differences in how disclosure affects individuals depending on whether they are already attached to a partner or searching for one. We therefore view baseline partnership status not only as a proxy for underlying quality, but also as capturing an economically important distinction in how criminal record disclosure shapes relationship opportunities, relationship stability, and downstream behavioral responses.

5.3.1 Family formation

Appendix Table A.6 reports results on the role of search duration, search market composition, and mobility among unpartnered men. Column (1) shows the effect of public disclosure on the number of years until finding a partner, conditional on eventually finding one. We find no significant effect on time to partnership, suggesting that the overall decline in partnership formation documented earlier reflects permanent exclusion from the marriage market rather than delayed matching.

To examine whether these effects are driven by changes in the composition of search markets, we proxy childhood social networks using indicators for whether partners attended the same high school or lived in the same neighborhood at age 16. We similarly proxy adult social networks by whether partners worked at the same workplace or lived in the same neighborhood three years prior to the launch of Lexbase. Conditional on eventually finding a partner, we find no significant effect of public disclosure on the likelihood of forming partnerships within either childhood or adult social networks. Taken together with survey evidence indicating that most individuals at the time met partners either online or through their social networks, these results suggest that listing did not substantially alter the venues through which individuals searched for partners.

Column (4) shows the effect of public disclosure on the probability of changing place of residence. We find no significant effect. This suggests that unpartnered men did not respond to public disclosure by relocating to a different residential area.

Table 4 provides evidence aimed at better understanding our finding that partnered men were less likely to remain partnered after public disclosure. Previous research (e.g. Aizer, 2010; Getik, 2024) suggests that relative bargaining power matters for family dynamics. Following Aizer (2010), we measure relative bargaining power within the household using

relative earnings and educational attainment (years of schooling). We then stratify the sample according to whether the female partner has higher earnings or more education than the male partner, which serves as a proxy for greater relative bargaining power. The results in columns (1) to (4) show that the effect of public disclosure on the probability of remaining partnered is stronger in couples where the female partner has relatively greater bargaining power.

We also stratify the sample by relationship length using the median duration in our sample. The results in columns (5) and (6) suggest that the effect of public disclosure on the probability of remaining partnered is concentrated among relatively new relationships. In these unions, it is less likely that the female partner had prior knowledge of the man's criminal record.

5.3.2 Violence against women

One possibility is that our results on violence against women simply reflect changes in reporting behavior, if public listing makes women more likely to report a crime conditional on a crime having actually occurred. The results in Appendix Table A.7 show no differential effect on violence against women for charges that resulted in conviction relative to charges that were dismissed, suggesting limited scope for changes in reporting behavior to explain the results.

Another possibility is that the effect on violence—and potentially also on family formation—is mediated through increased economic exclusion. However, the results in Appendix Table A.7 show no significant effects on men's own labor market outcomes.

Next, we ask to what extent changes in the characteristics of the new partners of previously unpartnered men can explain our findings. The results in Table 5 provide evidence of increased sorting toward women with a higher predicted risk of becoming victims of domestic violence. Specifically, using baseline characteristics, we predict the probability that a woman experiences domestic violence over a two-decade horizon and show that public listing increases the probability that a future partner belongs to the top quartile of the predicted risk distribution by 15.8 percentage points. At the same time, listing reduces the probability that a future partner has completed high school by 6.1 percentage points. These patterns are consistent with listed men disproportionately matching with more vulnerable partners.

It is also possible that public disclosure directly affects known correlates of IPV, such as mental health and substance use. Table 6 reports results for outcomes along these dimensions. The dependent variables in columns (1)–(3) are indicators for prescribed med-

ications in 2014–2018. The dependent variable in column (4) is an indicator for whether the individual was convicted of an alcohol- or drug-related offense (e.g. DUI or possession) during 2014–2018. We find significant increases in prescriptions for stress-related diagnoses, as well as for alcohol- and drug-related medications, among unpartnered listed men. These findings are consistent with criminological theories emphasizing that public stigmatization may exacerbate psychological distress and frustration, thereby increasing violent behavior.

5.4 Did the database protect women?

In Table 7, we use additional data to investigate the identity of female victims. These victims may be current or former partners, acquaintances, or strangers. The data allow us to distinguish whether the victim was a stranger to the offender. The dependent variable in column (2) is an indicator for whether the victim of the violent crime was a female stranger. The dependent variable in column (3) is an indicator for whether the victim was not a stranger (i.e. a partner, friend, co-worker, or classmate). The results suggest that the effect is concentrated among female victims who were not strangers to the offender.

Although we find that public disclosure increases the risk of violence against women, it is still possible that open-access crime databases allow women to avoid risky matches. We are unable to examine effects on women who were not already in a relationship, but we can conduct this analysis for women who were partnered with these men at baseline. While we find no statistically significant effect on violence against women in this group, the results in Table 8 show improvements in women's labor market outcomes. In particular, public listing increases women's probability of employment by 11.8 percentage points and reduces the probability of sickness absence by 11.7 percentage points. Taken together, these findings are consistent with public listing enabling women to exit troubled relationships without increasing their exposure to violence.

5.5 Role for public policy?

Appendix Table A.8 reports results by offenders' baseline mental health treatment status. The results show that the effects on violence against women are muted among men who had been treated for depression prior to listing, highlighting the role of psychological vulnerability and stress responses.

6 Concluding Remarks

We study the unexpected launch of Sweden's first online criminal records database in 2014, which provided unrestricted, anonymous, and cost-free public access to the criminal histories of all individuals charged with a crime. To estimate the causal effects of public access to criminal records, we combine rich population-wide administrative data with an institutional feature of the platform's design: administrative disclosure rules restricted the company behind the database from accessing records resulting from charges that occurred before specific cutoff dates. This constraint generates quasi-random variation in public exposure to criminal records, allowing us to isolate the informational effect of public listing from both underlying individual characteristics and formal criminal sanctions.

We document substantial heterogeneity by baseline partnership status. Among men who are partnered at baseline, public listing leads to a rapid 7.8 percentage points decline in the probability of remaining partnered, with stronger effects in newly formed relationships and in couples where the female partner has relatively greater bargaining power. We find no statistically significant effect on violence against women in this group, but observe improvements in women's labor market outcomes, consistent with public listing enabling women to exit troubled relationships without increasing exposure to violence. In contrast, among men who are unpartnered at baseline, public listing reduces the probability of entering a partnership by 8.7 percentage points and increases the risk of committing violence against women by 4.1 percentage points. These effects cannot be explained by changes in reporting behavior or adverse labor market outcomes. Instead, we document increases in prescriptions for stress-related diagnoses and alcohol or drug use, as well as evidence of increased sorting toward women with a higher predicted risk of becoming victims of domestic violence. The effects on violence against women are muted among men who were treated for depression prior to listing, highlighting the role of psychological vulnerability and stress responses.

Taken together, the results reveal a trade-off in the public disclosure of criminal records. While public access appears to improve match quality and facilitate exit from harmful relationships for partnered women, it may also exacerbate psychological distress and violent behavior among unpartnered listed men. These findings inform debates on criminal record transparency, public safety, and the unintended consequences of stigma-based policies.

References

Acquisti, A., Taylor, C., Wagman, L., 2016. The economics of privacy. Journal of Economic Literature 54, 442–492. URL: https://doi.org/10.1257/jel.54.2.442, doi:10.1257/jel.54.2.442.

Agan, A.Y., Starr, S.B., 2018. Ban the box, criminal records, and racial discrimination: A field experiment. The Quarterly Journal of Economics 133, 191–235. doi:10.1093/qje/qjx028.

Aizer, A., 2010. The gender wage gap and domestic violence. American Economic Review 100, 1847–1859. URL: https://doi.org/10.1257/aer.100.4.1847, doi:10.1257/aer.100.4.1847.

Akerlof, G.A., 1970. The market for "lemons": Quality uncertainty and the market mechanism. The Quarterly Journal of Economics 84, 488–500. URL: https://doi.org/10.2307/1879431, doi:10.2307/1879431.

Anderberg, D., Rainer, H., Wadsworth, J., Wilson, T., 2016. Unemployment and domestic violence: Theory and evidence. Economic Journal 126, 1947–1979.

Apel, R., 2013. Sanctions, perceptions, and crime: Implications for criminal deterrence. Journal of quantitative criminology 29, 67–101.

Avast, 2021. 50% of online dating services users research their date before their first face-to-face encounter. URL: https://www.prnewswire.com/news-releases/50-of-online-dating-services-users-research-their-date-before-their-first-face-to-html. survey of 15,233 Avast users from 15 countries about online dating safety and research behavior before first dates.

Backman, C., 2012. Criminal records in sweden: Regulation of access to criminal records and the use of criminal background checks by employers .

Becker, G.S., 1973. A theory of marriage: Part i. Journal of Political Economy 81, 813–846. doi:10.1086/260084.

Becker, G.S., 1974. A theory of marriage: Part ii. Journal of Political Economy 82, S11–S26. doi:10.1086/260170.

Becker, G.S., Landes, E.M., Michael, R.T., 1977. An economic analysis of marital instability. Journal of Political Economy 85, 1141–1187.

Bhuller, M., Dahl, G.B., Lø ken, K.V., Mogstad, M., 2024. Domestic violence reports and the mental health and well-being of victims and their children. Journal of Human Resources 59, S152–S186. URL: https://doi.org/10.3368/jhr.1222-12698R1, doi:10.3368/jhr.1222-12698R1. this study uses Norwegian administrative data to examine the effects of domestic violence reports on mental health and other outcomes for both victims and their children.

Bushway, S.D., Sweeten, G., 2007. Abolish lifetime bans for ex-felons. Criminology & Pub. Pol'y 6, 697.

Calabresi, E., Rodríguez-Planas, N., 2025. Female Empowerment and Intimate Partner Violence. IZA Discussion Paper 18207. IZA – Institute of Labor Economics. URL: https://www.iza.org/publications/dp/18207. discussion paper overviewing economic research on IPV and female empowerment.

Card, D., Dahl, G.B., 2011. Family violence and football: The effect of unexpected emotional cues on violent behavior. Quarterly Journal of Economics 126, 103–143.

Caucutt, E.M., Guner, N., Rauh, C., 2021. Incarceration, unemployment, and the racial marriage divide. BSE Working Paper 1300, November 2021. Available at: https://bse.eu/research/working-papers/racial-marriage-divide.

Cesarini, D., Lindqvist, E., Östling, R., Terskaya, A., 2023. Fortunate families? the effects of wealth on marriage and fertility. URL: http://www.nber.org/papers/w31039. nBER Working Paper No. 31039.

Charles, K.K., Luoh, M.C., 2010. Male incarceration, the marriage market, and female outcomes. The Review of Economics and Statistics 92, 614–627.

Charles, K.K., Stephens, M.J., 2004. Job displacement, disability, and divorce. Journal of Labor Economics 22, 489–522. doi:10.1086/381258.

Chiappori, P.A., Salanié, B., Weiss, Y., 2017. Partner choice, investment in children, and the marital college premium. Journal of Political Economy 125, 1791–1837.

Craigie, T.A., 2019. Ban the box, convictions, and public employment. Economic Inquiry 58, 425–445. doi:10.1111/ecin.12837.

Doleac, J.L., Hansen, B., 2018. The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden. Journal of Labor Economics URL: https://www.jenniferdoleac.com/wp-content/uploads/2018/09/Doleac_Hansen_JOLE_preprint.pdf. forthcoming; Texas A&M University Working Paper.

Evans, D.N., Begum, P., Porter, J.R., 2025. Love in the time of stigma: Public interest in relationships following criminal conviction. Journal of Crime and Justice 48, 96–112. doi:10.1080/0735648X.2024.2353236.

Evans, D.N., Vega, A., 2021. Experimental analysis of male online dating on parole. Journal of Crime and Justice 44, 33–48. doi:10.1080/0735648X.2020.1781678.

Fallesen, P., Andersen, L.H., 2017. Explaining the consequences of imprisonment for union formation and dissolution in denmark. Journal of Policy Analysis and Management 36, 154–177. doi:10.1002/pam.21933.

Getik, D., 2024. Relative income and mental health in couples. The Economic Journal 134, 3291–3305. URL: https://doi.org/10.1093/ej/ueae071, doi:10.1093/ej/ueae071, arXiv:https://academic.oup.com/ej/article-pdf/134/664/3291/60665174/ueae071.pdf.

Goldfarb, A., Tucker, C., 2019. Digital economics. Journal of Economic Literature 57, 3–43. URL: https://doi.org/10.1257/jel.20171452, doi:10.1257/jel.20171452.

Hankins, S., Hoekstra, M., 2011. Lucky in life, unlucky in love? the effect of random income shocks on marriage and divorce. Journal of Human Resources 46, 403–426.

Holzer, H.J., Raphael, S., Stoll, M.A., 2006a. Perceived criminality, criminal background checks, and the racial hiring practices of employers. Journal of Law and Economics 49, 451–480.

Holzer, H.J., Raphael, S., Stoll, M.A., 2006b. Perceived criminality, criminal background checks, and the racial hiring practices of employers. The Journal of Law and Economics 49, 451–480.

Jackson, O., Zhao, B., 2016. The effect of changing employers' access to criminal histories on ex-offenders' labor market outcomes: Evidence from the 2010–2012 massachusetts cori reform. URL: https://www.bostonfed.org/-/media/Documents/Workingpapers/PDF/wp1630.pdf. federal Reserve Bank of Boston Working Paper No. 16-30.

Kolesár, M., Rothe, C., 2018. Inference in regression discontinuity designs with a discrete running variable. American Economic Review 108, 2277–2304. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20160945, doi:10.1257/aer.20160945.

Liu, S., 2019. Incarceration of african american men and the impacts on women and children. URL: https://sitianliu.github.io/papers/siu_2019_no_appendix.pdf. unpublished Manuscript.

Lopoo, L.M., Western, B., 2005. Incarceration and the formation and stability of marital unions. Journal of Marriage and Family 67, 721–734. doi:10.1111/j.1741-3737.2005.00165.x.

NortonLifeLock, Poll, T.H., 2022. 2022 Cyber Safety Insights Report: Global Results. Technical Report. NortonLifeLock Inc. URL: https://www.nortonlifelock.com/content/dam/nortonlifelock/pdfs/reports/2022_NLCSIR_Global_Report.pdf. report based on an online survey of 10,003 adults in 10 countries, January 2022.

Ostrovsky, M., Schwarz, M., 2010. Information disclosure and unraveling in matching markets. American Economic Journal: Microeconomics 2, 34–63. URL: https://doi.org/10.1257/mic.2.2.34, doi:10.1257/mic.2.2.34.

Rose, E.K., 2021. Does banning the box help ex-offenders get jobs? evaluating the effects of a prominent example. Journal of Labor Economics 39, 79–113. doi:10.1086/711847.

Rosenfeld, M.J., Thomas, R.J., Hausen, S., 2019. Disintermediating your friends: How online dating in the united states displaces other ways of meeting. Proceedings of the National Academy of Sciences 116, 17753–17758. doi:10.1073/pnas.1908630116.

van Schellen, M., Poortman, A.R., Nieuwbeerta, P., 2012. Partners in crime? criminal offending, marriage formation, and partner selection. Journal of Research in Crime and Delinquency 49, 545–571. doi:10.1177/0022427811414197.

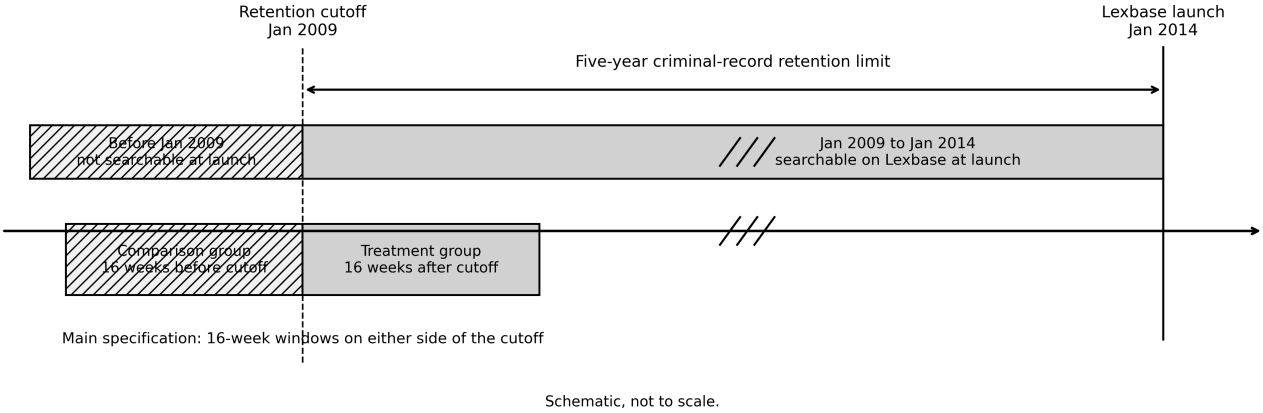
Siennick, S.E., Stewart, E.A., Staff, J., 2014. Explaining the association between incarceration and divorce. Criminology 52, 371–398.

Svarer, M., 2011. Crime and partnerships. Review of Economics of the Household 9, 307–325. doi:10.1007/s11150-010-9104-3.

Weiss, Y., Willis, R.J., 1997. Match quality, new information, and marital dissolution. Journal of Labor Economics 15, S293–S329. doi:10.1086/209864.

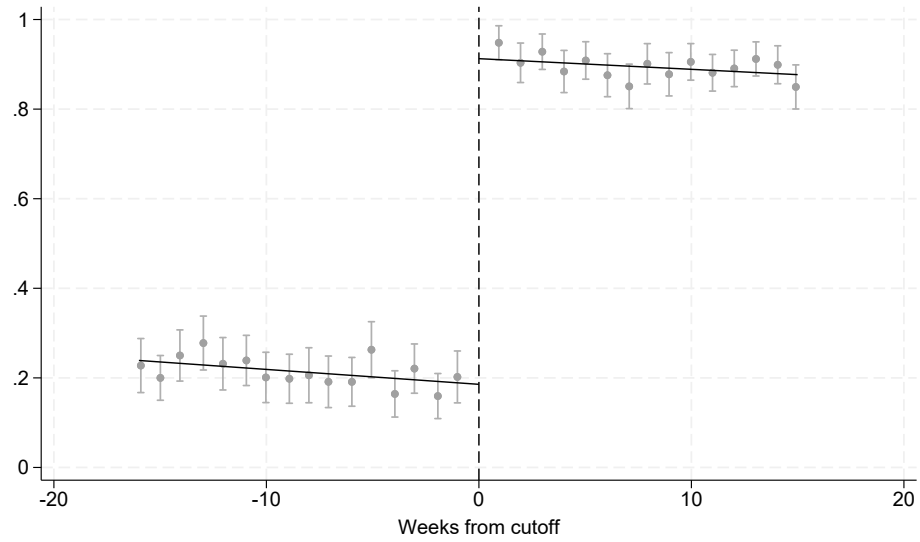
Figures and Tables

Figure 1: The retention period and treatment status



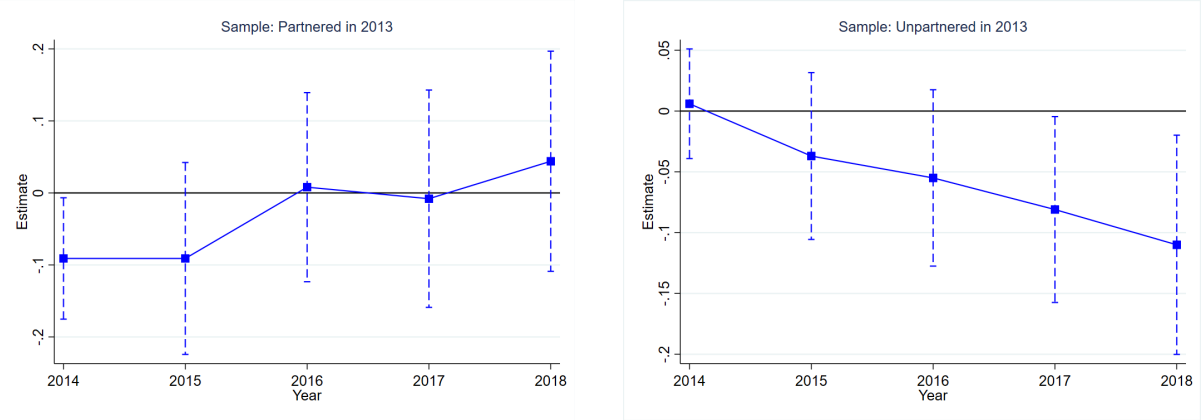
Note: Schematic (not to scale) illustration depicting the RD design for a hypothetical courthouse whose records were obtained in January 2014, when the Lexbase website was launched. With a five-year retention limit, the retention cutoff is January 2009. In our main specification, the treatment group consists of individuals who were first charged with a crime during the first 16 weeks of 2009. The comparison group consists of individuals first charged during the last 16 weeks of 2008. Outcomes are computed over the follow-up period extending from the January 2014 Lexbase launch date to December 2018.

Figure 2: First Stage



Note: The figure shows the probability of being publicly listed on the Lexbase launch date by the week in which the individual was first charged with a crime. The probability is greater than zero to the left of the cutoff because some individuals first charged prior to the cutoff are charged again after the cutoff. The probability is below unity after the cutoff because some courts delivered less than five full years of records, as discussed in section 4.3. The sample consists of first-time male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. The 95 percent confidence intervals are based on standard errors clustered at the court level. NEEDS A Y-AXIS LABEL THAT SAYS P(publicly listed).

Figure 3: IV estimates of the effect of public listing on the probability of having a partner by year



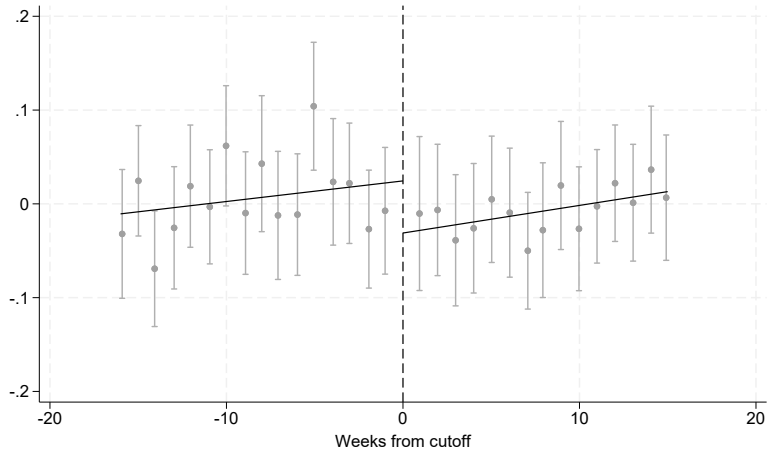
(a) With partner at baseline

(b) Without partner at baseline

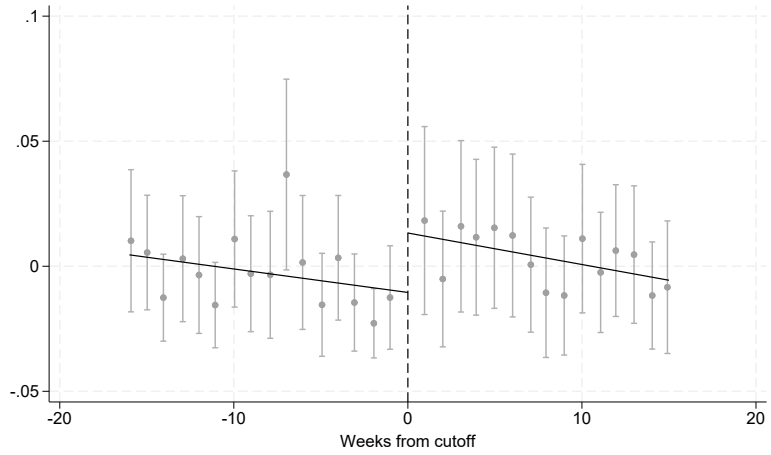
Notes: The sample consists of first-time male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. The dependent variable is an indicator for whether the individual had a partner in each year 2014-2018. Figure A shows year-by-year estimates for individuals with partners at baseline (1,561 observations) and Figure B shows year-by-year estimates for individuals without partners at baseline (4,598 observations). All regressions control for year and court fixed effects as well as for weeks-to-cutoff, allowing varying slopes above and below the cutoff. Baseline controls include the following covariates: age, immigrant, high school degree, college degree, missing education. Ninety-five percent confidence intervals are based on standard errors clustered at the court level. DROP TITLES ON THE STATA FIGURES.

Figure 4: Reduced form effect of being first charged after the cutoff

(a) Effect on the probability of having a partner



(b) Effect on the probability of committing violence against women



Notes: The sample consists of first-time male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. The dependent variable in Figure A is an indicator for whether the male has a partner in any year 2014-2018. The dependent variable in Figure B is an indicator for whether the male is convicted for OR CHARGED WITH? violence against women at least once in the years 2014-2018. Both figures plot the residulized outcomes after netting out year and court fixed effects and baseline controls: age, immigrant, high school degree, college degree, missing education. Ninety-five percent confidence intervals are based on standard errors that are clustered at the court level. I THOUGHT WE WERE GOING TO MOVE THIS TO THE APPENDIX? WE NEED Y-AXIS LABELS: (a) P(Having a partner), (b) P(Violence against women).

Table 1: Descriptive Statistics

	Publicly listed		First charged after the cutoff	
	Yes	No	Yes	No
Panel A: Background Characteristics	(1)	(2)	(3)	(4)
Foreign born	0.296	0.296	0.297	0.295
Age in 2013	31.012	33.372	31.855	32.349
Partnered in 2013	0.230	0.280	0.247	0.260
Compulsory degree in 2013	0.290	0.204	0.252	0.248
HS degree in 2013	0.518	0.561	0.536	0.541
College degree in 2013	0.175	0.217	0.194	0.196
Missing education	0.017	0.017	0.018	0.016
Panel B: Case characteristics				
Property conviction	0.132	0.101	0.113	0.121
Violent conviction	0.237	0.201	0.226	0.216
Drug conviction	0.055	0.048	0.052	0.051
DUI conviction	0.073	0.085	0.077	0.080
Traffic conviction	0.136	0.177	0.149	0.161
Other conviction	0.256	0.268	0.260	0.262
Acquitted	0.111	0.120	0.122	0.109
Prison sentence	0.067	0.072	0.070	0.069
Fines	0.272	0.250	0.266	0.258
Probation sentence	0.122	0.131	0.132	0.122
Panel C: Outcomes				
Has partner	0.390	0.451	0.406	0.429
Violence against women	0.077	0.037	0.064	0.053
Obs.	3,272	2,888	2,877	3,283

Note: The table shows mean values for our analysis sample consisting of male defendants aged 20 to 60 who first charged within a 16-week window around the court-specific cutoff. Public listing is the endogenous variable and First charged after the cutoff is the instrument.

Table 2: First-stage, reduced form and 2SLS estimates of being listed on the probability of having a partner, by baseline partnership status

	Mean	OLS	2SLS	2SLS
Panel A: Full sample	(1)	(2)	(3)	(4)
Has partner	0.419	-0.069***	-0.098*	-0.090**
	(0.493)	(0.020)	(0.050)	(0.043)
Panel B: With partner at baseline				
Has partner	0.955	-0.022*	-0.080**	-0.078**
	(0.208)	(0.012)	(0.035)	(0.035)
Panel C: Without partner at baseline				
Has partner	0.236	-0.029*	-0.085*	-0.087*
	(0.425)	(0.015)	(0.047)	(0.047)
Court & FE	–	Yes	Yes	Yes
Baseline Controls	–	No	No	Yes
Obs.		6,160	6,160	6,160
Obs. partnered		1,562	1,562	1,562
Obs. unpartnered		4,598	4,598	4,598

Note: The table shows estimates of the effect of being listed on the probability of having a partner at any time during 2014-2018. Panel A-C show separate estimates for different subgroups. The sample consists of male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. All regressions control for year and court fixed effects as well as for weeks-to-cutoff, allowing variying slopes above and below the cutoff. Baseline controls include age, immigrant status, high school degree, college degree, missing education. The first-stage coefficients in columns (2)-(5) are .733 (.033), .732 (.033), .777 (.051) and .719 (.033). The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Table 3: Effect of public listing on violence against women

	Mean	OLS	2SLS	2SLS
Panel A: Full sample	(1)	(2)	(3)	(4)
Violence against women	0.036	0.028***	0.039**	0.038**
	(0.187)	(0.006)	(0.019)	(0.018)
Panel B: With partner at baseline				
Violence against women	0.031	0.013	0.030	0.028
	(0.174)	(0.012)	(0.030)	(0.030)
Panel C: Without partner at basEline				
Violence against women	0.038	0.032***	0.041*	0.041*
	(0.191)	(0.007)	(0.024)	(0.024)
Court & FE	–	Yes	Yes	Yes
Baseline Controls	–	No	No	Yes
Obs.		6,160	6,160	6,160
Obs. partnered		1,562	1,562	1,562
Obs. unpartnered		4,598	4,598	4,598

Note: The table shows estimates of the effect of public listing on the probability of committing violence against women at any time during 2014-2018. Panels A-C show separate estimates for different subgroups. The sample consists of male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. All regressions control for year and court fixed effects as well as for weeks-to-cutoff, allowing variying slopes above and below the cutoff. Baseline controls include age, immigrant status, high school degree, college degree, missing education. The first-stage coefficients in columns (2)-(5) ARE .823 (.047), .824 (.046), .894 (.053) and .782 (.057). The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Table 4: Effect of public listing on probability of having a partner among men with partners at baseline, by partner and relationship characteristics

	Partner has Higher education	Partner has Lower education	Partner has Higher earnings	Partner has Lower Earnings	Short Relationship	Long Relationship
	(1)	(2)	(3)	(4)	(5)	(6)
Partnered	-0.134***	-0.050	-0.119	-0.056*	-0.115**	-0.034
	(0.049)	(0.039)	(0.082)	(0.032)	(0.046)	(0.053)
	[0.955]	[0.954]	[0.957]	[0.953]	[0.940]	[0.972]
Court & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	627	935	445	1,117	856	706

Note: The table shows the effect of public listing on the probability of having a partner IN EACH YEAR? during 2014-2018. The sample consists of male defendants aged 20 to 60 who had partners at baseline and who were first charged within a 16-week window around the court-specific cutoff. All regressions control for year and court fixed effects and age, immigrant status, high school degree, college degree, missing education. The cutoff for the sample split on relationship length is 6 years, which is the median length in our data. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level. NOT SURE I'M RIGHT ABOUT THE DEP. VAR. BUT IF IT IS A DUMMY FOR EVER HAVING A PARTNER DURING 2014-18, I DON'T SEE HOW THE EFFECTS COULD BE SO LARGE, CONSIDERING THAT THESE GUYS STARTED 2014 WITH A PARTNER

Table 5: Characteristics of new partners among men without partners at baseline

	Education above High school	Employment	Earnings	Conviction in Past 10 years	High predicted risk of Becoming victim of IPV
	(1)	(2)	(3)	(4)	(5)
	-0.061***	0.133	-0.009	0.002	0.158**
	(0.023)	(0.089)	(0.204)	(0.070)	(0.070)
	[0.079]	[0.729]	[12.059]	[0.116]	[0.261]
Court & Year FE	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Obs.	4,598	4,598	4,598	4,598	4,598

Note: The sample consists of male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff, who were without partners at baseline, and who found a partner during 2014-2018. The outcomes pertain to the new partners. All regressions control for year and court fixed effects and age, immigrant status, high school degree, college degree, missing education. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level. SAMPLE SIZES CAN'T BE RIGHT.

Table 6: Effect of public listing on mental health, alcohol, and drug use among men without partners at baseline

	Any prescription	Anti-depressants	Sleep/anxiety	Alcohol/drug arrest
	(1)	(2)	(3)	(4)
	0.015	0.025	0.080*	0.060
	(0.035)	(0.045)	(0.048)	(0.037)
	[0.831]	[0.201]	[0.241]	[0.188]
Court & Year FE	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes
Obs.	4,598	4,598	4,598	4,598

Note: The baseline sample consists of male defendants aged 20 to 60 who where sentenced within a 16-week window around the court-specific cutoff and were without partners at baseline (WHY THIS RESTRICTION?) . The dependent variables in columns (1)-(3) are indicators for prescribed medications in 2014-2018. The dependent variable in column (4) is an indicator for whether the individual has been convicted for an alcohol or drug related crime (e.g. DUI or possession) in 2014-2018. All regressions control for year and court fixed effects and age, immigrant status, high school degree, college degree, missing education. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level. I THINK WE SHOULD DROP THIS. THERE IS ONE MARGINALLY SIGNIFICANT COEFFICIENT AMONG FOUR. EITHER THAT OR ADD THE GUYS WITH PARTNERS AND SEE IF THE ESTIMATES BECOME MORE PRECISE

Table 7: Effect of public listing on violence among men without partners at baseline, by whether victim was known or stranger

	Baseline	Stranger to The victim	Known To the victim
	(1)	(2)	(3)
	0.041*	0.011	0.058**
	(0.024)	(0.009)	(0.026)
	[.]	[0.009]	[0.059]
Court & Year FE	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes
Obs.	4,598	4,598	4,598

Note: The sample consists of male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff and who were without partners at baseline. The dependent variable in column (2) is an indicator for whether the defendant committed violence against a female stranger. The dependent variable in column (3) is an indicator for whether he committed violence against a woman who was not a stranger (i.e. partners, friends, co-workers, classmates). All regressions control for year and court fixed effects and age, immigrant, high school degree, college degree, missing education. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

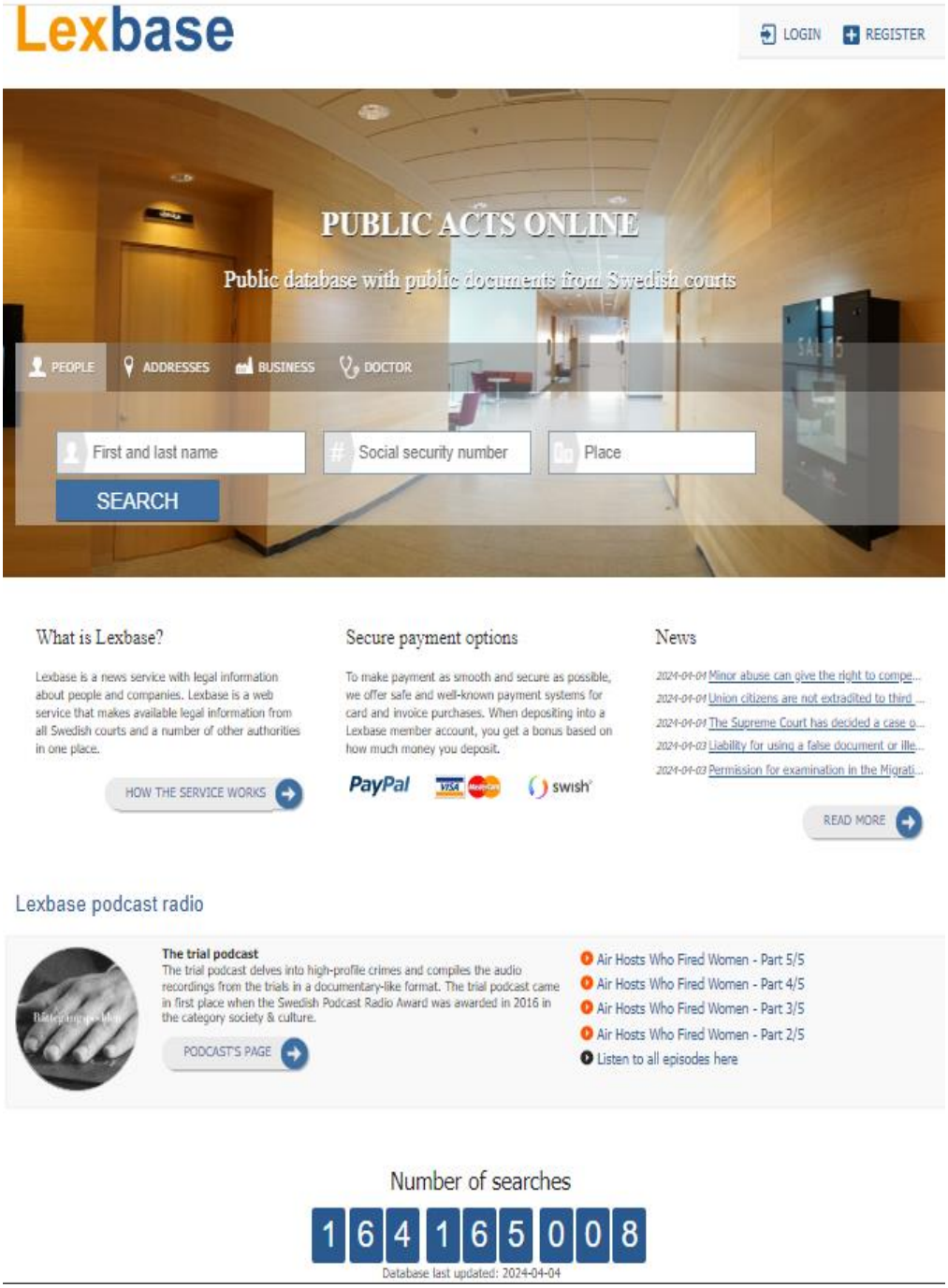
Table 8: Effects of public listing on the outcomes of the partners of males who had partners at baseline

	Employed	Earnings	Sickness absence	Anti-depressants	Sleep/anxiety	Victim
	(1)	(2)	(3)	(4)	(5)	(6)
	0.118*	0.109	-0.119**	-0.035	0.098	0.001
	(0.068)	(0.175)	(0.060)	(0.066)	(0.067)	(0.006)
	[0.650]	[12.126]	[0.388]	[0.250]	[0.241]	[0.006]
Court & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,562	1,562	1,562	1,562	1,562	1,562

Note: The sample consists of women who in 2013 where partners of male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. Earnings is log average annual earnings in 2014-2018. Victim is an indicator for having received in- or out-patient care for domestic violence in 2014-2018. All regressions control for year and court fixed effects and age, immigrant, high school degree, college degree, missing education. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Appendix

Figure A.1: Lexbase Homepage



Notes: This image shows the homepage of Lexbase and was retrieved on April 4, 2024.

How to use Lexbase

With a few single button presses, you can obtain useful decision-making information. You start by choosing a search method. The guide below explains how the service works.

1. Select search method

A) Search person
In this search function Search person you can specify your search using four fields; First name, last name, social security number and city. If you get a hit on a person, it means that he or she has been subject to a court hearing. If you do not get a hit, it is probably because the person does not appear in any legal process.

Example: Shows results for Anders Andersson in Stockholm

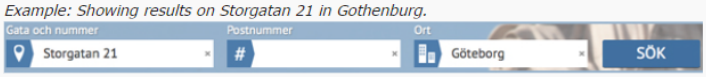


In the search fields, you can use asterisk (*) in your search text to replace part of a search term. You also don't need to search for an exact name of a person or company. Parts of the name in question are sufficient.

If you do not get a hit on a certain person, it is probably because they are not present in any legal process.

B) Search address
This search function helps you who want information about people living in a certain area or in your immediate area. The search tool sorts by Address, Postcode and City. The more fields you fill in, the narrower the hit picture you get.

Example: Showing results on Storgatan 21 in Gothenburg.



2. Search results are displayed

When you have made a search, the search result is presented in a list. To read a report, click on the button on the right which looks like this and is called FÖREKOMMANDE

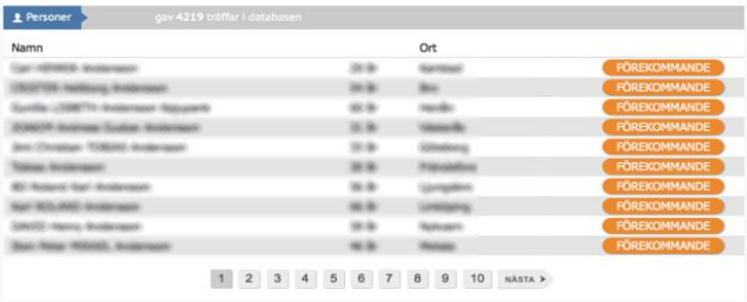
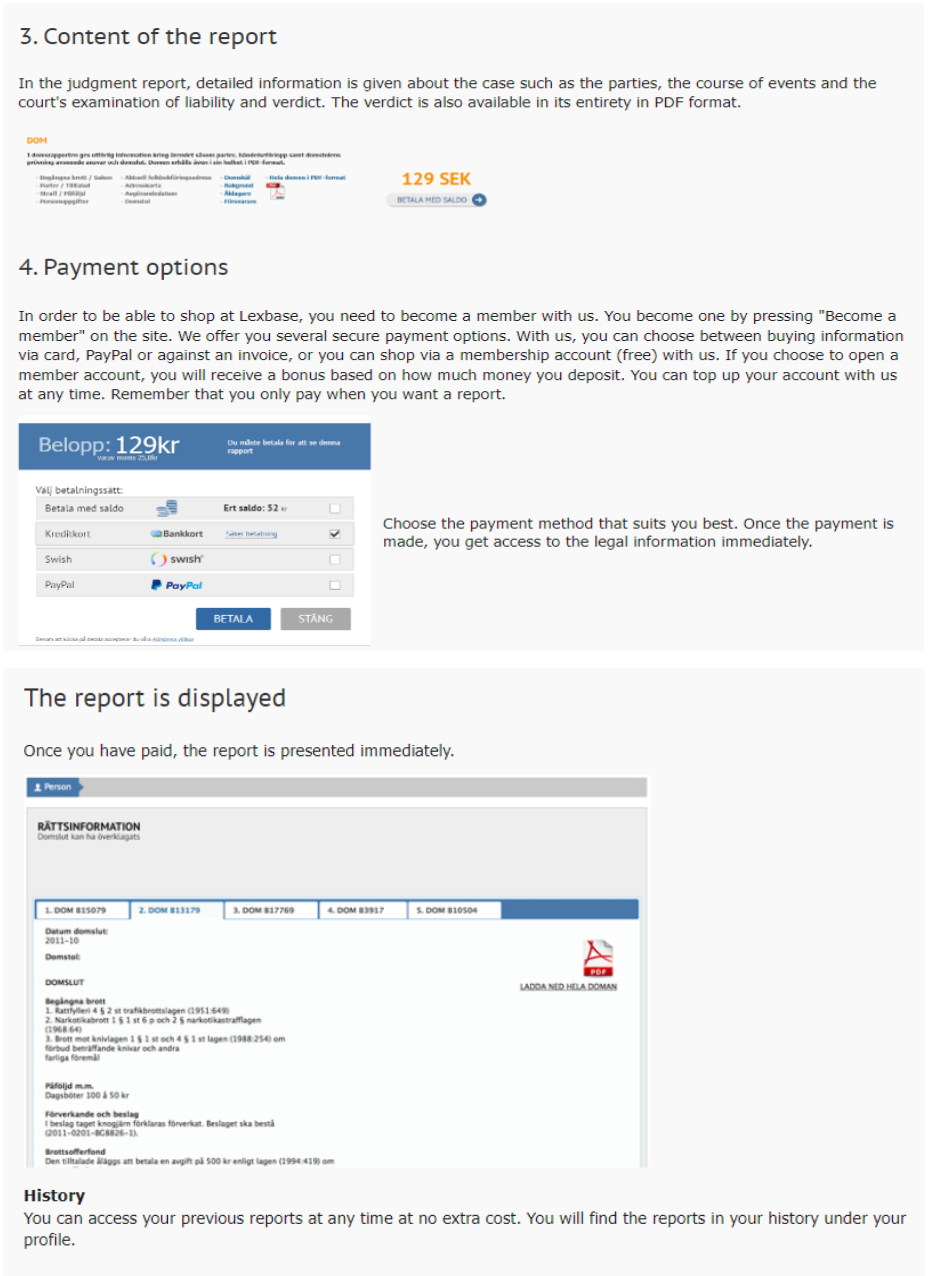


Figure A.2: How the Search Engine at Lexbase Works



Notes: This image is from the homepage of Lexbase and provides the guide from the platform on how to use the search engine. The image was retrieved on April 4, 2024.

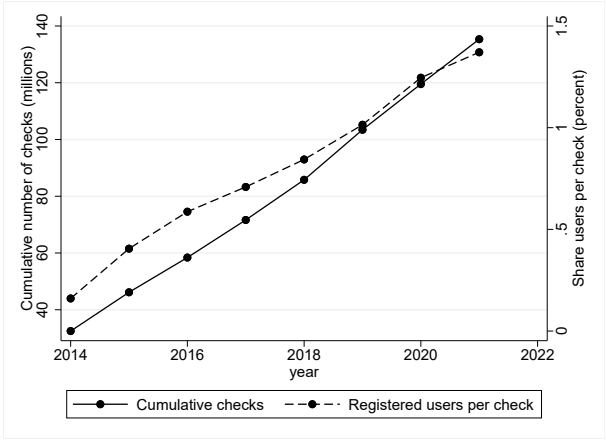


Figure A.3: Background checks using Lexbase

Notes: This figure plots the cumulative number of searches for criminal background using the platform. The figure also plots the number of registered users per background check conducted. The figure uses data provided by Lexbase. WHAT IS ADDED BY THE NUMBER OF REGISTERED USERS PER BACKGROUND CHECK?

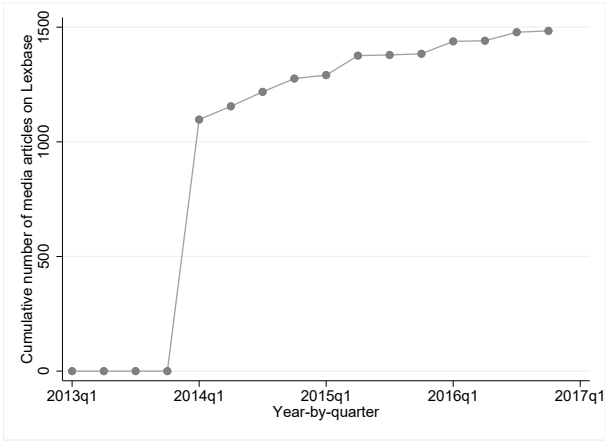
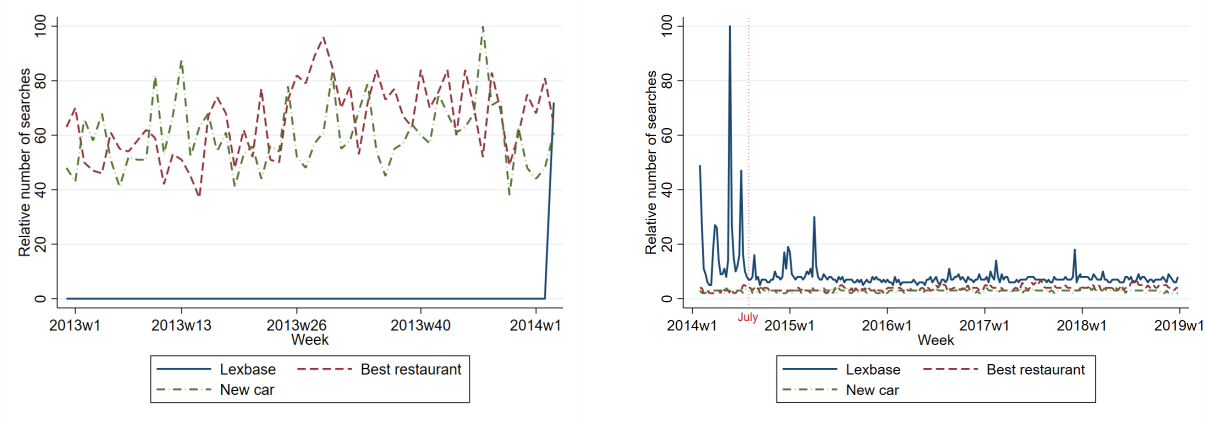


Figure A.4: Cumulative number of media articles covering Lexbase, 2013-2016

Notes: This figure plots the cumulative number of media articles mentioning Lexbase in 2013-2016. Data on media coverage is from the Retriever/Mediearkivet database, which is a Nordic news archive with material from mainly major Swedish newspapers. The information was collected on Jan 20th, 2024.



(a) Ending one week before launch

(b) Starting one week after launch

Figure A.5: Google searches in Sweden. Searches for Lexbase were 48–60% higher than for those for "Best restaurant" and "New car" in the post period starting July 2014 JANUARY? WHY DON'T WE JUST DO ONE GRAPH FROM JAN 2013 TO 2018?.

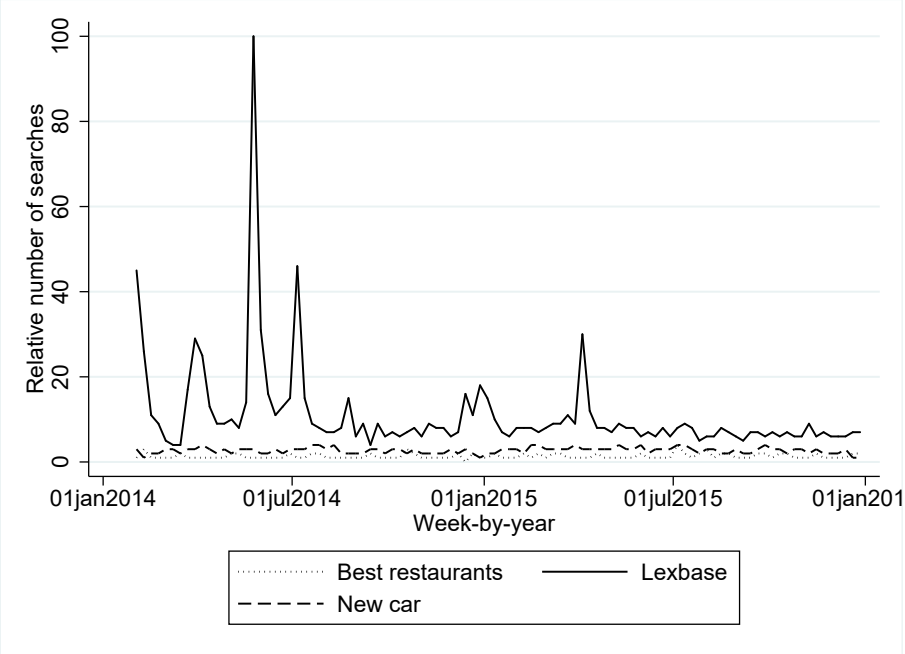


Figure A.6: Relative number of searches on Google

Notes: This figure plots the relative number of Google searches for the terms "Lexbase," "Best restaurants," and "New car." The data were collected from Google Trends on January 20, 2024. Google Trends normalizes search data to facilitate comparisons across terms. Search intensity is normalized by dividing each observation by the total number of searches in the relevant geography and time period, and then scaling the resulting values from 0 to 100. To improve readability, we omit the days immediately following the launch of the database, when searches spiked sharply, and instead focus on the period beginning one week after the introduction and ending in the last week of 2015.

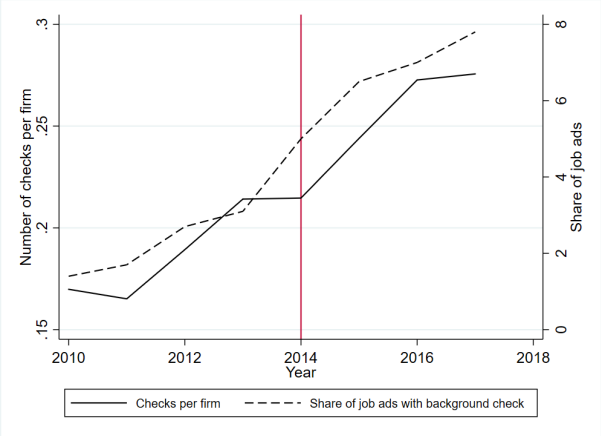
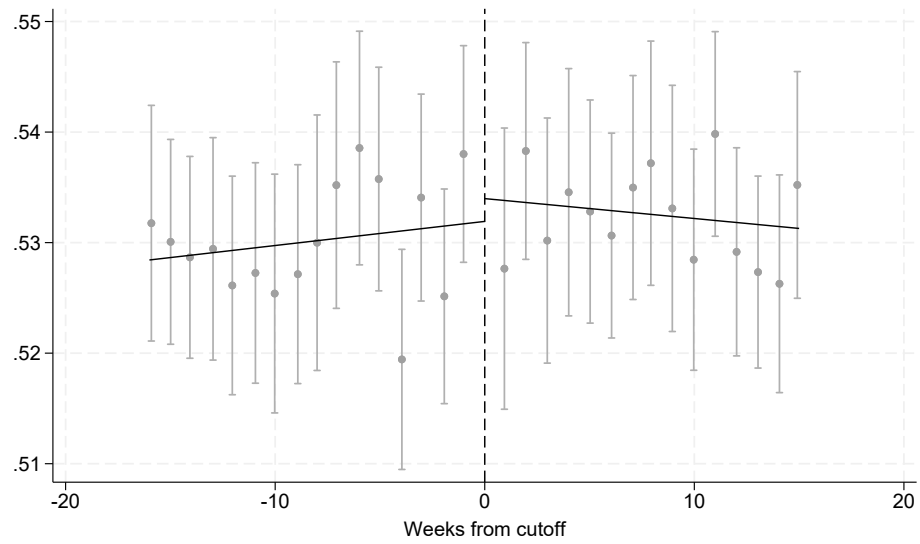


Figure A.7: Official (non-Lexbase) criminal background checks

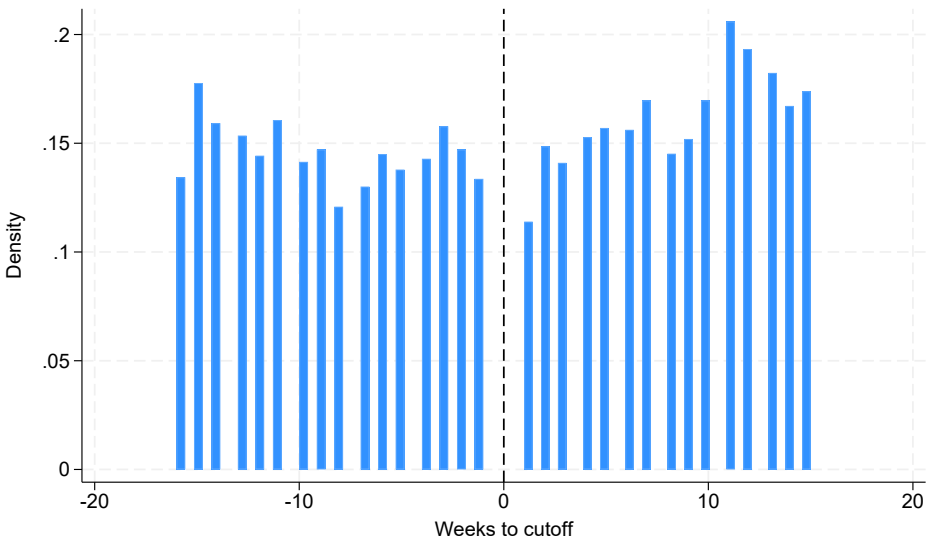
Notes: This figure plots the number of official criminal background check inquires for personal use (likely to be submitted to a potential employer) per registered firm submitted to the Swedish police over the period 2010-2017 using data from the Swedish police and Statistics Sweden. The figure also plots the share of ads posted at the Public Employment Office that mentions criminal background checks using data from the Public Employment Office. THIS IS REALLY CONFUSING. I DON'T KNOW WHO IS REQUESTING OR SUBMITTING WHAT TO WHOM.

Figure A.8: Balance plot



Note: The sample consists of first-time male defendant aged 20 to 60 who where sentenced in a 16-week window around the court-specific cutoff. The dependent variable is the predicted value of the endogenous variable "publicly listed" from a regression where this variable is regressed on baseline covariates. The prediction is plotted against weeks until the court-specific cutoff. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level. I DON'T UNDERSTAND THIS. WHY WOULD THE PREDICTED VALUE OF THE TREATMENT VARIABLE BALANCE ACROSS THE CUTOFF WINDOW?

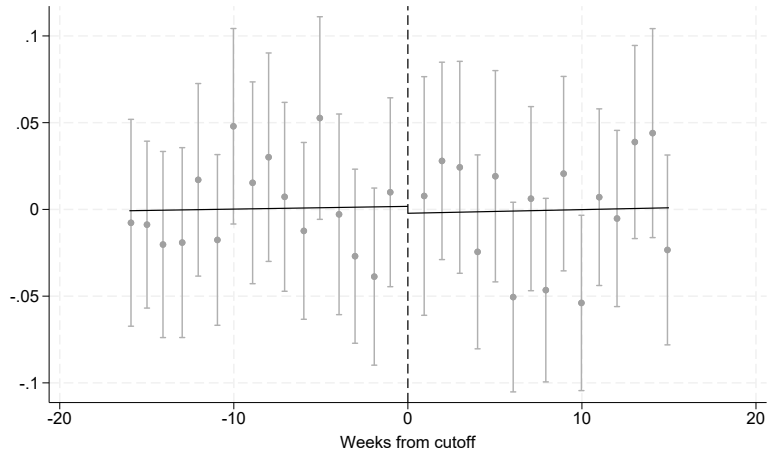
Figure A.9: Density of first-time charges around the cutoff



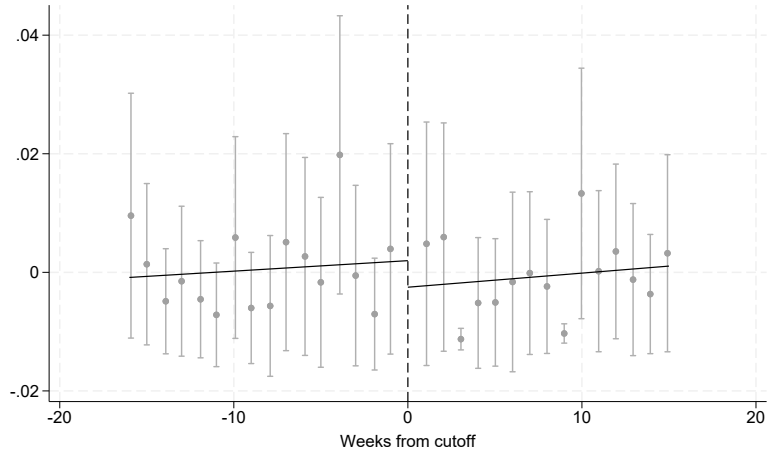
Note: The sample consists of male defendant aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff at week 0. RDD estimates show no significant change in the (log) number of cases around the cutoff (est. 0.141; se 0.207). THE SPACING ON THIS FIGURE IS ODD. IT SEEMS LIKE IT SHOULD BE UNIFORM, BUT IT IS NOT. IS THERE A WAY TO MAKE IT MORESO? ALSO, I THINK THIS WOULD LOOK NICER IF THE BARS GRAY AND WINDER, SO THEY TOUCHED ONE ANOTHER. THE NUMBER OF BARS TO THE LEFT OF ZERO IS 16, WHICH SEEMS RIGHT, BUT THERE ARE ONLY 15 TO THE RIGHT. SHOULDN'T THERE BE A BAR AT ZERO? FIGURE NEEDS A Y-AXIS LABEL: P(First charged).

Figure A.10: Reduced form plots for placebo year 2013

(a) Effect of first being charged after the cutoff on the probability of having a partner in 2013



(b) Effect of first being charged after the cutoff on the probability of committing violence against women in 2013



Notes: The sample consists of male defendants aged 20 to 60 who first charged within a 16-week window around the court-specific cutoff. The dependent variable in Figure A is an indicator for whether the male has a partner in 2013, before Lexbase. The dependent variable in Figure B is an indicator for whether the male is convicted of (CHARGED WITH?) violence against women in 2013. Both figures plot the residulized outcomes after netting out year and court fixed effects and age, immigrant status, high school degree, college degree, missing education. Ninety-five percent confidence intervals are based on standard errors that are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Table A.1: First Stage Results by Subgroups

	Sample Mean	Sentenced after the cutoff
	(1)	(2)
Full sample	0.521	0.703***
	(0.500)	(0.028)
Foreign Born	0.532	0.654***
	(0.499)	(0.035)
Native Born	0.531	0.692***
	(0.499)	(0.028)
Aged 20-30	0.572	0.637***
	(0.495)	(0.026)
Aged 31-40	0.496	0.733***
	(0.500)	(0.039)
Aged 41-50	0.467	0.768***
	(0.499)	(0.042)
Aged 51-60	0.452	0.745***
	(0.498)	(0.046)
Compulsory degree	0.617	0.545***
	(0.486)	(0.029)
HS degree	0.511	0.704***
	(0.500)	(0.032)
College degree	0.477	0.810***
	(0.500)	(0.033)
Court & Year FE	–	Yes
Baseline Controls	–	Yes

Note: The sample consists of male defendants aged 20 to 60 who who were first charged within a 16-week window around the court-specific cutoff. The table shows the first-stage estimate for subgroups stratified by values of the covariates in the baseline model. All regressions control for the other covariates as well as year and court fixed effects and weeks-to-cutoff, allowing for variying slopes above and below the cutoff. The Sanderson-Windmeijer F-statistic in the full sample is 654.9 and the Cragg-Donald F-statistic of 4569.7 The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Table A.2: Balancing

	Publicly listed	First charged after the cutoff
	(1)	(2)
Age in 2013	-0.003***	-0.001**
	(0.001)	(0.001)
Foreign born	0.019	0.011
	(0.013)	(0.011)
HS degree in 2013	-0.081***	0.001
	(0.015)	(0.012)
College degree in 2013	-0.087***	0.008
	(0.016)	(0.016)
Missing education	-0.076	0.011
	(0.049)	(0.044)
Violent conviction	0.058**	-0.013
	(0.024)	(0.023)
Property conviction	0.067**	-0.044
	(0.034)	(0.031)
Traffic conviction	-0.010	-0.044**
	(0.024)	(0.020)
Drug conviction	0.039	-0.055*
	(0.034)	(0.031)
DUI conviction	0.025	-0.018
	(0.030)	(0.028)
Prison sentence	0.012	0.037*
	(0.028)	(0.021)
Fines	0.011	0.016
	(0.020)	(0.015)
Probation sentence	-0.030	0.023
	(0.020)	(0.019)
Joint F-Test	[0.00000]	[0.23399]
Observations	6,160	6,160

Note: The sample consists of male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. Column (1) shows results from a regression of an indicator for being publicly listed (endogenous variable) on all baseline covariates. Column (2) shows results from a joint regression of an indicator for first being charged after the cutoff (instrument) on all baseline covariates. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Table A.3: Effect of public listing on the probability of having a partner among females

	Mean	OLS	2SLS	2SLS
Panel A: Full sample	(1)	(2)	(3)	(4)
Has partner	0.459	-0.056**	0.005	0.005
	(0.498)	(0.026)	(0.055)	(0.054)
Panel B: With partner at baseline				
Has partner	0.935	-0.005	0.006	0.007
	(0.246)	(0.019)	(0.057)	(0.056)
Panel C: Without partner at basline				
Has partner	0.240	-0.021	0.088	0.088
	(0.427)	(0.024)	(0.058)	(0.057)
Court & FE	–	Yes	Yes	Yes
Baseline Controls	–	No	No	No
Obs.		2,061	2,061	2,061
Obs. partnered		648	648	648
Obs. unpartnered		1,413	1,413	1,413

Note: The sample consists of female defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. The dependent variable is an indicator for whether the individual had a partner in any year 2014-2018. All regressions control for year and court fixed effects as well as for weeks-to-cutoff, allowing varying slopes above and below the cutoff. Controls include age, immigrant status, high school degree, college degree, missing education. The first-stage coefficient in column (2)-(5) is .823 (.047), .824 (.046), .894 (.053) and .782 (.057). The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Table A.4: Robustness and additional results

	Discrete RD	Bandwidth 20 weeks	Bandwidth 8 weeks	Placebo Year	Multiple time Offenders	Misdemeanor	Felony	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Has partner	-0.106***	-0.085***	-0.015	-0.018	-0.041	-0.153**	-0.069	
	(0.039)		(0.032)	(0.052)	(0.031)	(0.030)	(0.064)	(0.054)
Violence against women	0.034**	0.027**	0.053**	-0.008	-0.022	0.001	0.058***	
	(0.016)		(0.013)	(0.026)	(0.009)	(0.021)	(0.017)	(0.022)
Court & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Obs.	6,160	7,717	2,800	6,160	17,852	2,471	3,689	

Note: The baseline sample consists of male defendants aged 20 to 60 who were first charged within a 16-week window around the court-specific cutoff. All regressions control for year and court fixed effects as well as age, immigrant status, high school degree, college degree, missing education. Row (1) controls for the running variable using year-by-month FEs. Row (2) increases the bandwidth to 20 weeks. Row (2) shrinks the bandwidth to 8 weeks and estimates a local randomization RD without controlling for the running variable. Row (4) uses as a placebo the dependent variable measured in the year before the launch of Lexbase. Row (5) includes also multiple time offenders in the sample. I DON'T UNDERSTAND THIS. I THOUGHT WE ALWAYS INCLUDED PEOPLE WHO OFFENDED AGAIN AFTER THE CUTOFF. Row (6) and (7) stratify the sample based on whether the the average incarceration rate for the first charge was below 1 percent (Misdemeanor) or above 1 percent (Felony). PERCENT? The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Table A.5: Sample characteristics for men with and without partners at baseline

	With partners	Without partners
Panel A: Demographic Characteristics	(1)	(2)
Foreign born	0.426	0.252
Age in 2013	39.736	29.531
Panel B: Socioeconomic characteristics		
Compulsory degree in 2013	0.196	0.268
HS degree in 2013	0.498	0.552
College degree in 2013	0.291	0.162
Missing education	0.015	0.018
Employment in 2013	0.808	0.614
Earnings in 2013	346235.551	200201.835
Panel B: Crime		
Post conviction	0.147	0.281
Post prison sentence	0.025	0.065
Post violent conviction	0.031	0.090
Post number of convictions	0.259	0.598
Panel B: Health		
On anti-depressants	0.069	0.082
On sleep/anxiety medication	0.067	0.089
On ADHD medication	0.006	0.023
Obs.	1,562	4,598

Note: The baseline sample consists of male defendants aged 20 to 60 who were first charge withing a 16-week window around the court-specific cutoff. Convictions and incarceration are measured from the retention cutoff to the launch of the website.

Table A.6: Search duration, search market composition, and mobility for men without partners at baseline

	Time to find partner	Partner from childhood social network	Partner from adulthood social network	Change place of residence
	(1)	(2)	(3)	(4)
	-0.052	-0.000	0.038	-0.058
	(0.341)	(0.076)	(0.115)	(0.048)
	[2.870]	[0.170]	[0.339]	[0.638]
Court & Year FE	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes
Obs.	4,598	1,087	1,087	4,598

Note: The sample consists of male defendants aged 20 to 60 without partners at baseline who first charged within a 16-week window around the court-specific cutoff. The sample in column (1)-(3) is conditional on having found a partner in five years after Lexbase. The dependent variable in column (1) is years until finding a partner. The dependent variable in column (2) is an an indicator for the partner being from same high school or neighborhood (at age 16). The dependent variable in column (3) is an indicator for the partner being from the same workplace or neighborhood (three years before Lexbase). All regressions control for year and court fixed effects as well as age, immigrant, high school degree, college degree, missing education. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level. I DON'T THINK THESE SUBSAMPLE DESCRIPTIONS ARE CORRECT

Table A.7: Changes in reporting behavior and labor market outcomes

	Baseline	Case Charged	Effect on Employment	Effect on Earnings
	(1)	(2)	(3)	(4)
	0.041*	0.045*	0.011	-0.255
	(0.024)	(0.025)	(0.038)	(0.242)
	[0.038]	[0.036]	[0.853]	[11.733]
Court & Year FE	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes
Obs.	4,598	4,598	4,598	4,598

Note: The sample consists of male defendants aged 20 to 60 without partners at baseline who were first charged within a 16-week window around the court-specific cutoff. The dependent variable in column (2) is an indicator set to unity if the prosecutor following the investigation assessed the evidence as sufficiently strong to bring it to trial and zero otherwise. The dependent variable in column (3) is an indicator for being employed in any year 2014-2018. The dependent variable in column (4) is (log) average earnings in 2014-2018 with zero earnings replaced with 1. All regressions control for year and court fixed effects as well as age, immigrant, high school degree, college degree, missing education. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

Table A.8: Effects of public listing on XXX by offender's mental health treatment in 2013

	Baseline	Anti-depressant in 2013	No anti-depressant in 2013
	(1)	(2)	(3)
	0.041*	-0.027	0.074***
	(0.024)	(0.058)	(0.026)
	[0.912]	[0.937]	[0.916]
Court & Year FE	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes
Obs.	4,598	191	1,082

Note: The baseline sample consists of male defendants aged 20 to 60 who were first charged withing a 16-week window around the court-specific cutoff. DEPENDENT VARIABLE? All regressions control for year and court fixed effects aw well as age, immigrant, high school degree, college degree, missing education. The standard errors are clustered at the court level. */**/*** denote significance at the 10/5/1 percent level.

‘

A Online Appendix

Interpreting the IV Estimand Under Treatment Effect Heterogeneity

This subsection provides a simple framework for interpreting the IV estimand in our setting and for understanding why the IV estimates may exceed the corresponding OLS estimates in absolute value. To keep notation simple, we abstract from never-takers generated by imperfect matching or incomplete recovery of Lexbase records; allowing for such cases would not change the main intuition.

Let $W_i \in \{0, 1\}$ indicate whether defendant i's original case falls within the Lexbase disclosure window. Let $R_i \in \{0, 1\}$ indicate whether defendant i commits a subsequent offense before the launch of Lexbase that would itself place him in the database. We assume that near the cutoff, W_i is as-good-as-random and independent of R_i and of potential outcomes.

In the simplest sharp version of the model, listing status at the time of Lexbase's launch is given by

$$D_i(W_i) = \mathbb{1}\{W_i = 1 \text{ or } R_i = 1\}. \tag{3}$$

Thus, defendants with $R_i = 1$ are *always-takers*: they appear on Lexbase regardless of whether the original case falls inside the disclosure window. Defendants with $R_i = 0$ are *compliers*: they appear on Lexbase if and only if the original case falls within the disclosure window. In this formulation, compliers are precisely those defendants whose public visibility is determined by the age of the original case.

Let potential outcomes satisfy

$$Y_i(d) = \alpha + \eta R_i + [\tau_C(1 - R_i) + \tau_A R_i]\, d + \varepsilon_i, \tag{4}$$

where τ_C is the treatment effect for compliers, τ_A is the treatment effect for always-takers, and η captures type-specific differences in untreated outcomes. We assume $\mathbb{E}[\varepsilon_i \mid W_i, R_i] = 0$.

Let $\pi \equiv \Pr(R_i = 1)$. Then the first stage is

$$\mathbb{E}[D_i \mid W_i = 1] - \mathbb{E}[D_i \mid W_i = 0] = 1 - \pi, \tag{5}$$

since only defendants with $R_i = 0$ change listing status when W_i switches from 0 to 1. The reduced form is

$$\mathbb{E}[Y_i \mid W_i = 1] - \mathbb{E}[Y_i \mid W_i = 0] = (1 - \pi)\tau_C. \tag{6}$$

It follows that the Wald estimand is

$$\beta_{IV} = \frac{\mathbb{E}[Y_i \mid W_i = 1] - \mathbb{E}[Y_i \mid W_i = 0]}{\mathbb{E}[D_i \mid W_i = 1] - \mathbb{E}[D_i \mid W_i = 0]} = \tau_C. \tag{7}$$

The IV estimator therefore identifies the effect of public disclosure for defendants whose visibility is changed by the disclosure cutoff—that is, for defendants whose earlier case would otherwise have remained practically obscure.

By contrast, an OLS comparison by realized listing status pools together compliers and always-takers. Let $q \equiv \Pr(W_i = 1)$. Then

$$\beta_{OLS} = \mathbb{E}[Y_i \mid D_i = 1] - \mathbb{E}[Y_i \mid D_i = 0] = \omega\tau_C + (1-\omega)\tau_A + (1-\omega)\eta, \tag{8}$$

where

$$\omega = \Pr(R_i = 0 \mid D_i = 1) = \frac{q(1-\pi)}{q(1-\pi)+\pi}. \tag{9}$$

Hence, after conditioning on rich covariates so that differences in untreated outcomes are limited (that is, $\eta \approx 0$), OLS is a weighted average of the treatment effect for compliers and the treatment effect for always-takers. If $|\tau_C| > |\tau_A|$ and the two effects have the same sign, then

$$|\beta_{OLS}| < |\beta_{IV}| = |\tau_C|. \tag{10}$$

The intuition is straightforward. For compliers, Lexbase makes an otherwise aging criminal record newly searchable, so the informational shock is large. For always-takers, by contrast, a subsequent offense would have made the defendant visible anyway, so the marginal informational content of the older case is smaller. In that case, the IV estimator isolates the group for whom disclosure matters most, while OLS averages this effect together with the smaller effect for offenders who would have appeared in the database regardless.

When $\eta \neq 0$, OLS also reflects selection on recidivism risk, since always-takers may differ from compliers in untreated outcomes. This selection term can in principle move OLS either toward or away from the IV estimate. For that reason, the interpretation that IV exceeds OLS because treatment effects are stronger for compliers is most compelling when the comparison is made within the local RD sample and with rich controls. An extension allowing for misclassification in measured Lexbase appearance would attenuate OLS further and therefore reinforces the possibility that IV exceeds OLS in absolute value.

B Interpreting our LATE

This section includes additional details on how we calculate the number and characteristics always takers, never takers, and compliers in our sample.

Overview: Following Dahl et al. (2014), we define compliers as children whose parents' incarceration decision would have been different had their case been assigned to the most strict instead of the least strict judge:

$$\pi_c = Pr(Prison_i = 1|Z_i = \overline{z}) - Pr(Prison_i = 1|Z_i = \underline{z}) = Pr(Prison_i(\overline{z}) > Prison_i(\underline{z}))$$

where π_c represents the fraction of compliers, $\overline{z}$ represents the maximum value of our judge instrument (the most stringent judge) and $\underline{z}$ represents the minimum value of our instrument (the least stringent judge).

Always takers are children whose parents would always be incarcerated regardless of the judge assigned to their case. Because of the monotonicity and independence assumptions, the fraction of always takers is given by the probability of a parent being incarcerated by the least stringent judge:

$$\pi_a = Pr(Prison_i = 1|Z_i = \underline{z}) = Pr(Prison_i(\overline{z}) = Prison_i(\underline{z}) = 1)$$

Finally, never takers are children whose parents would never be incarcerated, with the fraction of never takers given by the probability of not being incarcerated by the most stringent judge:

$$\pi_n = Pr(Prison_i = 0|Z_i = \overline{z}) = Pr(Prison_i(\overline{z}) = Prison_i(\underline{z}) = 0)$$

Number of Compliers: We calculate the shares of children in each category by looking at the rates of parental incarceration for the "most lenient" and "most strict" judges. In Table A.1, we estimate our linear specification of the first stage of incarceration on our residualized measure of judge stringency controlling for exhaustive court-by-time fixed effects, under different definitions of the most lenient and most stringent judges. Under our preferred specification where we define most lenient judge as the bottom 1 percentile of judge stringency and the most strict judge as the top 1 percentile of judge stringency, we find that 17 percent of our sample are compliers, 67 percent are never takers, and 16 percent are always takers.

Characteristics of Compliers: We recover the characteristics of our complier population

by calculating the fraction of compliers in different subsamples (Abadie 2003; Dahl et al. 2014). We find that compliers are much more likely to be charged with a drug offense, but much less likely to be charged with a violent offense. They are somewhat more likely to have a prior conviction and to be unemployed at baseline.

Appendix B References

.15in1

Abadie, Alberto. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics*, 113(2): 231-263.

Dahl, Gordon B., Andreas Ravndal Kostøl, and Magne Mogstad. 2014. "Family Welfare Cultures."*Quarterly Journal of Economics*, 129(4): 1711-1752.

Table A.1: Sample Share by Compliance Type

	1 percent	1.5 percent	2 percent
Compliers	0.149	0.137	0.122
Never Takers	0.674	0.677	0.683
Always Takers	0.177	0.186	0.195

Note: This table reports the the share of always takers, never takers, and compliers in our sample. Compliers are defined as children whose parents' incarceration decision would have been different had their case been assigned to the most lenient instead of the most strict judge. Always takers are instead parents who would never be released regardless of the judge assigned to their case. Finally, never takers are parents who would never be incarcerated regardless of the stringency of the judge. Following Dahl et al. (2014) we calculate the shares of parents in each category by looking at the prison rates for parents assigned to the most lenient and most strict judges. We define the most lenient judge as the bottom 1/1.5/2 percentile of judge stringency and the most strict judge as the top 1/1.5/2 percentile of judge stringency. We estimate our linear specification of the first stage to recover compliers as the share of parents predicted to get incarcerated at the top percentile minus the share of parents predicted to get incarcerated at the bottom percentile, always takers as the share of parents predicted to be sentenced to prison at the bottom percentile and never takers as the share of parents who are predicted to be released at the top percentile.

Table A.2: Characteristics of Marginal Defendants

	$P[X=x]$	$P[X=x\|complier]$	$\frac{P[X=x\|complier]}{P[X=x]}$
Child is native born	0.914	0.916	1.002
	(0.001)	(0.015)	(0.016)
Child is foreign born	0.086	0.084	0.976
	(0.001)	(0.015)	(0.172)
Low education parent	0.803	0.797	0.993
	(0.002)	(0.022)	(0.027)
High education parent	0.197	0.203	1.027
	(0.002)	(0.022)	(0.109)
Parent with prior conviction	0.495	0.639	1.292
	(0.002)	(0.028)	(0.057)
Parent with no prior conviction	0.505	0.361	0.714
	(0.002)	(0.028)	(0.056)
Parent employed at baseline	0.280	0.145	0.520
	(0.002)	(0.028)	(0.100)
Parent not employed at baseline	0.720	0.855	1.186
	(0.002)	(0.028)	(0.039)
Parent charged with property offense	0.150	0.153	1.024
	(0.001)	(0.021)	(0.139)
Parent charged with non-property offense	0.850	0.847	0.996
	(0.001)	(0.021)	(0.024)
Parent charged with violent offense	0.140	0.038	0.270
	(0.001)	(0.027)	(0.189)
Parent charged with non-violence offense	0.860	0.962	1.119
	(0.001)	(0.027)	(0.031)
Parent charged with drug offense	0.062	0.138	2.223
	(0.001)	(0.017)	(0.264)
Parent charged with non-drug offense	0.938	0.862	0.919
	(0.001)	(0.017)	(0.018)
Parent charged with DUI offense	0.105	0.120	1.135
	(0.001)	(0.021)	(0.196)
Parent charged with non-DUI offense	0.895	0.880	0.984
	(0.001)	(0.021)	(0.023)

Note: This table presents the sample distribution, complier distribution, and relative likelihood for different subgroups. Bootstrapped standard errors in parentheses are obtained using 500 replications.

C Data Appendix

Incarceration: An indicator for whether a defendant is incarcerated in any year 2014-2018.

Male: An indicator for whether the defendant is male.

Native Born: An indicator equal to one if the defendant was born in Sweden.

Criminal Conviction An indicator for whether a defendant is convicted of any crime 2014 to 2018.

Property Conviction An indicator for whether a defendant is convicted of property crime 2014 to 2018.

Violent Conviction An indicator for whether a defendant is convicted of violent crime 2014 to 2018.

Drug Conviction An indicator for whether a defendant is convicted of a drug crime between 2014 and 2018.

High School Degree or Above An indicator for whether a defendant has a high school degree or a higher education degree (e.g., college, university) in 2013.

Employment An indicator for whether the defendant is employed in 2013. Employment is measured as having positive earnings in a given year.

Earnings ($1,000s) Nominal values are deflated to 2015 and represented in U.S. dollars using the exchange rate SEK/$ = 9.25.

Number of Children, Parent: Counts the number of children the charged parent has in the three years preceding their trial date.

Welfare Use i Measures the welfare receipt status of the charged employed in the 5 years following their trial date. Calculated as the average of 5 indicators (one for each of years

1-5 post-Lexbase). Each indicator is set equal to one if the household of the charged parent receives welfare payments in a given year.

With partner This proxies whether the charged parent is married or cohabiting with common children.

Without partner This proxies whether the charged parent is not married and cohabiting with common children.