School and Crime

Todd R. Jones and Ezra Karger*

September 16, 2024

Abstract

Criminal activity is seasonal, peaking in the summer and declining through the winter. We provide the first evidence that arrests of children and reported crimes involving children follow a different pattern: peaking during the school year and declining in the summer. We use a regression discontinuity design surrounding school start dates and an excess crime calculation to show that the school environment increases reported crimes involving children by roughly 50% annually. School exacerbates preexisting inequality in criminal interactions, increasing the Black-white and male-female gaps in reported juvenile crime and arrest rates by more than 40%.

*Jones: Mississippi State University, CESifo, and IZA. toddrjones.com. Karger: Federal Reserve Bank of Chicago. ezrakarger.com. Any views expressed do not necessarily reflect those of the Federal Reserve Bank of Chicago or the Federal Reserve System. For useful feedback, we would like to thank seminar participants at the 2022 NBER Spring Children's Meetings, the 2022 NBER Spring Education Meetings, the 2022 Southeastern Micro Labor Conference, the 2022 Southern Economic Association Meetings, University of Connecticut, and University of South Carolina, as well as Daniel Aaronson, Robert Ainsworth, Eunsik Chang, Will Davis, Jeffrey Denning, Jennifer Doleac, David Figlio, Joshua Goodman, Oded Gurantz, Brian Jacob, Kendall Kennedy, Lars Lefgren, Bhashkar Mazumder, Valerie Michelman, Steve Ross, and Jeff Swigert. We would also like to thank Abbie Natkin, Rosa Kleinman, and Haoyang Xiong for research assistance.

1 Introduction

Reported rates of criminal activity are highest for teenagers, declining sharply through age 20 (Shulman et al., 2013). While many papers document this relationship between age and criminal activity, there is minimal research into the causes or predictors of early-age interactions with the criminal justice system. We use the exact timing of reported criminal offenses to show that school plays a major causal role in driving higher rates of reported and realized criminal activity among children. Our data do not allow us to disentangle this causal effect into components due to different features of the school environment. But, our results highlight the importance of school-based law enforcement interactions as key drivers of the lifelong relationship between children and the carceral state.

A large literature investigates seasonal patterns in crime rates. Block (1984) documents seasonality in crime by crime type, highlighting 180 years of academic work in a similar vein dating back to Quetelet (1842). Quetelet showed that person-on-person crime peaked in the summer and dropped in the winter. Recent evidence from a panel of U.S. cities finds modern-day empirical support for these patterns–with overall criminal activity peaking in the summer (McDowall et al., 2012). We use incident-level data from the National Incident-Based Reporting System (NIBRS) to trace out seasonality in criminal activity by age, confirming the general finding that overall criminal activity peaks in the summer.[1] But, we then focus on children aged 10–17 and show that this group does not follow the overall pattern: criminal activity involving children as offenders, victims, and arrestees peaks during the school year (especially in the fall and spring) and falls during the summer.[2]

To show that this descriptive fact reflects a causal relationship, we use the start and/or end dates of the school year for over 3,000 school districts linked to nearby law enforce-

[1]See Appendix Figure A.1 for all reported crimes (Panel A) and all arrests (Panel B) using 2017-2019 NIBRS data.

[2]In two papers that closely parallel our work, Hansen and Lang (2011) and Hansen et al. (2022) show that suicide rates jump when school is in session.

ment agencies—and a regression discontinuity design—to show that this reversed pattern of seasonality is largely caused by the school environment, which leads to a dramatic increase in rates of juvenile arrests (41%) and reported crime (47%). We draw on the work of Fitzpatrick et al. (2020), who use a similar method to show that the start of the school year leads to a sharp increase in reports of child maltreatment.[3] We also build on work from Jacob and Lefgren (2003), who use teacher-in-service days to study short-term breaks from school and find that being in school causes lower levels of property crime but higher levels of violent crime among children. Our work relates to Luallen (2006), who show that teacher strikes lead to higher rates of property crime but lower rates of violent crime; Akee et al. (2014), who show that time away from school decreases crime; Fischer and Argyle (2018), who shows that there is higher crime (particularly property crime) when schools change to a four-day school week; and Bacher-Hicks et al. (2022), who show that in-person school closures resulting from the Covid-19 pandemic decreased school bullying and cyberbullying. While these papers support our causal story—that schools cause increases in crime involving children—they focus on specific policies and quasi-experimental variation that may not generalize to the nationwide seasonal trend in crime rates involving children that we document in this paper.

We use data from the National Crime Victimization Survey to show that our results are not driven by differential reporting of crimes in the summer and during the school year–we find the same patterns described above when we analyze monthly rates of victimization self-reported by a nationally-representative sample of respondents in The National Crime Victimization Survey (NCVS). This result is mirrored by prior work from Carbone-Lopez and Lauritsen (2013) who are the first we know of to descriptively document a different pattern of seasonality in victimization rates among children and adults (albeit at a monthly level, and focused on violent crime rates), also using the NCVS. We additionally use the NCVS to show similar patterns for offenders.

[3]In a related paper, Cowan et al. (2023) use school calendar information to study how teenagers and parents of school-age children change their time use between the school year and summer.

We explore heterogeneity in this seasonal pattern of crime for 10–17-year-olds and show that the increase in crime during the school year is driven by crime reported during school hours (7am to 8pm) on weekdays, crimes involving acquaintances, drug crimes, assaults, and intimidation (but not property crimes), and crimes reported as occurring in school. In fact, we see the typical (for adults) seasonal pattern of crime peaking in the summer for crimes involving 10–17-year-olds that are reported as having occurred in a location outside of school.

To better understand the exact geography of these crimes, we focus on two cities—Austin and New York City—which provide additional information about reported crimes. In Austin, TX, there are two distinct law enforcement agencies that report crimes separately: the Austin Police Department and the Austin Independent School District Police Department. The Austin Police Department shows the standard unimodal seasonality in reported crimes involving 10–17-year-olds—with the number of crimes peaking in the summer—while the Austin Independent School District Police Department shows strong patterns of seasonality consistent with the national trends described above for crimes involving children. Relatedly, in New York City we observe the latitude and longitude of reported crimes, and we show that the bimodal seasonal patterns described above are explained by reported crimes within 0.25 miles of a school; and we see the strongest patterns for crimes reported from significantly closer to the school. We see little evidence of bimodal seasonal patterns aligned with the school year for crimes involving 10–17-year-olds that are reported more than 0.25 miles away from a school.

We conclude our paper by quantifying the fraction of 'excess' crime among 10–17-year-olds that is caused by the school environment. We granularly control for general seasonal patterns in reported crime during non-school hours and we calculate residual criminal incidents among each age group in our sample, building on a typical framework for calculating excess mortality (Wang et al., 2022). We find little residual ('excess') crime for older cohorts (aged 19+) during the school year and school day, but we find strong evidence of school-related seasonality for crimes reported involving 10–17-year-olds. On the average day, 10–17-year-olds are involved in around 270 additional crimes and 70 additional arrests in our sample of law enforcement agencies reporting to NIBRS. Within our sample, we argue that schools cause a 47% increase in reported crimes and a 41% increase in reported arrests among 10–17-year-olds annually. We then focus on demographic subgroups and argue that the school environment (separately) causes an approximately 47% increase in reported crimes for male, female, Black, and white children. These roughly 47% increases magnify gaps in reported crimes by demographic group, increasing annual Black-white and male-female gaps in reported crime rates by roughly 47% relative to the gaps we would expect in a world with no 'excess' reported crime related to the school environment. We find little evidence that excess crime varies as a function of student demographics, school resources, or other characteristics of local communities. In other words, the pattern we find is remarkably universal: in poor and rich counties; well-resourced school districts and poorly resourced school districts; and rural and urban counties, schools are a primary driver of criminal activity involving children.

Our results contrast with a large literature measuring the causal effect of educational attainment on criminal activity. Recent work by Bell et al. (2022) uses state-level variation in dropout laws to show that laws preventing children from dropping out of high school caused a dynamic form of educational incapacitation, reducing criminal activity by keeping at-risk children in school. That work builds on related research by Anderson (2014), Machin et al. (2011), and Hjalmarsson et al. (2015) who use sharp variation in school dropout laws and compulsory schooling laws to show that additional years of completed education cause reductions in criminal activity. While this literature might seem at odds with our results, that is not necessarily the case. Prior work measures the causal effect of education on criminal activity for the set of compliers who respond to compulsory schooling laws and dropout laws in various contexts. That set of compliers—those students on the margin of not completing an additional year of schooling until a law caused them to stay in school for an extra year—represent a small minority of the students enrolled in school at any given point in time.[4]

We show that in equilibrium, across all students, school causes large increases in crimes involving (and arrests of) children relative to a counterfactual where crime involving children follows typical seasonal patterns. Our results in no way suggest that school fails a cost-benefit test: there are many other measured benefits that students receive from school. But, the contrast between our findings and prior work on compulsory schooling laws does caution against extrapolating from specific policy's local average treatment effects (eg. compulsory schooling laws reduce crime) to a claim about the costs and benefits of more general policies and institutions, like the claim that in aggregate, schools reduce crime: a claim our results contradict.

2 Empirical strategy

2.1 Data and Sample Selection

Our primary data source for reported crime is the National Incident-Based Reporting System, or NIBRS, and we focus on the years 2017-2019.[5] NIBRS is a crime reporting system used by many, but not all, police departments across the United States. In 2019, agencies reporting to NIBRS covered around half of the U.S. population, documenting criminal incidents in areas with 147 million U.S. residents and 8,500 law enforcement agencies.[6] NIBRS contains information on the exact date and hour of each incident; the type of incident; the age, race, and sex of the victims and offenders; and the relationships

[4]This must be the case because school enrollment before the implementation of these laws is often 80-90% depending on the age group, country, and context. And the students enrolled in school prior to the law change were necessarily not compliers to the law.

[5]We obtain these data from Kaplan (2021a).

[6]See `https://www.fbi.gov/news/pressrel/press-releases/fbi-releases-2019-nibrs-crime-data` for more details.

between the victims and offenders. Not all variables are observed for every incident. In addition, we observe if the incident resulted in an arrest or citation.[7] In most, but not all, cases, the arrests we consider are a subset of the incidents we consider. Each incident corresponds to a law enforcement agency (police agency).[8]. We construct the main sample such that there is one observation per incident, even if the incident involved multiple offenders. We associate each incident with the youngest offender in the incident (considering only offenders at least 10 years old). This leads to younger ages being over-represented to some degree, though most incidents involve only one offender. We also restrict to police agencies that reported in all 12 months of the year.

We conduct our analysis at the weekly level. To do so, we assign to each incident the NIBRS incident date (though in a minority of cases cases, this is the reported date) We aggregate the number of reported crimes or arrests to the weekly level.[9] NIBRS reports information for more than one type of offense and for multiple offenders and victims. Throughout, we consider only the first-listed offense type and information for only one of the offenders and one of the victims victim.[10]

For much of our analysis, we use only the NIBRS data described above, referred to as the 'main sample.' However, we also explore a regression discontinuity design (RDD), using school start and end dates to demonstrate a causal link between reported crime and school. We refer to this sample, described below, as the 'RDD sample.'

To implement our RDD, we first gather school calendar information for the 2019–2020 school year from a calendar that aggregates school holiday, start date, and end date in-

[7] In particular, we observe if the incident resulted in the offender being 1) "taken into custody (based on warrant and/or previous incident report)"; 2) given an "on-view arrest (taken into custody without a warrant or previous incident report)"; or 3) "summoned/cited (not taken into custody)." We include categories 1 and 2 when constructing our arrest variable.

[8] We do not refer to a law enforcement agency as an LEA to avoid confusion with a local education agency, which shares the same acronym.

[9] We focus on calendar weeks of the year such that the first Monday of the year is in the first week and the next-occurring Sunday is the final day of the first week. We assign weeks such that the first seven days are week 1, the next seven days are week 2, etc. We repeat until week 52, which includes 8 days.

[10] See the Data Appendix for more details about this and other datasets.

formation for public consumption: publicholidays.com.[11] The school calendar data initially includes 11,245 public school districts districts, and for nearly every district we observe both the first and last days of school (as well as Thanksgiving, Christmas, and spring breaks).[12] We make use of a subset of these districts (roughly 2,500), after matching the school calendar data to the NIBRS data. Because our school calendar data contain only state and school district name, in an intermediate step we obtain the school district FIPS county code after using fuzzy matching to the National Center for Education Statistics (NCES) Education Demographic and Geographic Estimates Public School District file from 2019–2020 (National Center for Education Statistics, 2020). We then link NIBRS incidents to the school start (end) week. For our RDD analysis, to reduce measurement error we keep only incidents from police agencies in counties for which all districts in the sample have the same start (end) week. This has the advantage of being conservative, but the disadvantage of dropping many observations from the data (that correspond to multiple school districts with differing start (end) dates). We present further details on how we constructed this dataset in the Data Appendix. Our final RDD sample for the start of school (before any additional sample restrictions, such as requiring a valid age of offender) consists of observations from 1,270 police agencies representing 2,560 school districts. For the end of school, there are 1,147 police agencies representing 2,019 school districts.[13]

In order to provide the cleanest estimate, in our RDD analysis, we focus only on 2019 NIBRS linked to the 2019-2020 school calendar information. This means that the school start date is measured more accurately than the school end date because the incidents occurred during 2019, but the school calendar represents the school end date for 2020.[14]

[11]We also have data for 2020–2021 and 2021-2022, but focus on 2019-2020 to avoid confounding our estimates with the Covid-19 pandemic.

[12]We also observe holidays such as fall break, Easter break, and midwinter break, though these are less common.

[13]Combined, there are 1,412 police agencies representing 3,034 school districts that appear in at least one of the two samples.

[14]It is unclear if publicholidays.com updated the 2019-2020 school calendar to reflect any changes in the wake of the Covid-19 pandemic, but the results are generally sharp, leading us to believe that they did not.

However, it is our sense that school calendars do not change much year-to-year, and our results for school end date are generally sharp, suggesting this is not a cause for concern.[15]

In addition to the NIBRS data, we also analyze data from the 1992-2019 National Crime Victimization Survey (NCVS), a nationally representative survey measuring the number and types of victims of crimes in the U.S. (Bureau of Justice Statistics., 2021). By comparing seasonality for children's criminal behavior in NIBRS and the NCVS, we can explore whether the patterns we focus on are explained by changes in the reporting of crimes surrounding the school year. We find no evidence that our results (for crimes involving victims) are explained by reporting bias.

One shortcoming of NIBRS is that it does not include sub-police agency geocoded information. While NIBRS does include information on crime location (home, business, school, etc.), we do not observe the latitude and longitude of the reported crime. To fill this gap, we use New York City (NYC) incident-level crime data (New York City Police Department, 2022)[16] These data include the age range of the offender and victim, the date of incident, and (approximate) geocoded location of incident. We link each under 18-year-old offender incident to the nearest school, where the school location is obtained from the NCES Education Demographic and Geographic Estimates using the Public School file and Private School file from 2019–2020 (National Center for Education Statistics, 2020). We use this distance variable to consider how crime changes over the year for reported crimes various distances away from the nearest school.

2.2 Specification

Our main results rely on three types of analyses–first, because the results we present below are visually sharp, we focus on raw counts of reported crimes committed by of-

[15]There are of course other sources of measurement error. For example, while the school calendar data include many districts, they do not contain all.

[16]We considered 41 cities that had crime data and found only two—New York City and Cincinnati—that had age (range) of offender, date of incident, and geocoded location. Three others have age (range) of victim, date of incident, and geocoded location. We focus on New York City.

fenders in different demographic groups over different time periods. These plots of raw outcomes provide convincing evidence of the empirical patterns we study in this paper.

Second, to more precisely estimate our statistical claims, we rely on a regression discontinuity design to estimate the causal effect of children attending school on crime. We present plots in the style of a regression discontinuity design where we stack the set of crimes in even-time surrounding the week when schools first began (or finished) each calendar years. These plots show even sharper empirical results, but are largely consistent with the raw data.

Finally, we estimate the amount of 'excess crime' caused by school over the entire school year relative to a counterfactual where crime rates involving children follow typical adult patterns of seasonality. We detail this approach in Section 3.6.

3 Results

We begin by presenting several graphs that show seasonal patterns of reported crime in the 2017–2019 NIBRS data. First, we plot the total number of reported crimes by week in Figure 1. We limit the sample to offenders 10–17 years old in Panels A and B, to offenders 19–24 years old[17] in Panels C and D, and to offenders 25–30 years old in Panels E and F. The left Panels (A, C, and E) show all reported incidents while the right Panels (B, D, and F) subset to incidents resulting in the arrest of an offender.[18] Across each of these figures, we see an increase (or no change) in reported crime and arrests between the start of the year and roughly week 20 (occurring in May), and a decrease in reported crime from around week 35 to 40 (in September) to the end of the year. However, these patterns for each age group sharply diverge between calendar weeks 20 and 40 (May and Septem-

[17]We do not consider age 18 because it is an interstitial group, in which some children remain in school at this age while many others do not. However, the figures in this table look similar if we include 18-year-olds in the 10–17 group.

[18]For arrests, we rely on the age of the arrestee. We also use the date of the incident the arrest is associated with rather than the arrest date because the arrest date might be after the date of the incident and we are interested in when the actual crime occurs.

ber) for 10–17 year-olds relative to the other age groups. We see a dramatic decrease in reported crime and arrests for 10–17-year-olds beginning in May, and a sharp increase in reported crimes and arrests beginning at the end of the summer.[19][20] We also consider the victim's (as opposed to offender's) age in Appendix Figure A.3 and find similar results. Taken together, these findings suggest that school, which typically recesses during the summer, plays a role in the sharp divergence between the calendar week and reported crimes and arrests for 10–17-year-olds relative to other groups.

3.1 Regression Discontinuity Analysis

To establish whether this relationship is causal, we now consider patterns in reported crime by week relative to the estimated school district start date in Figure 2. These figures are limited to the RDD sample described above, with the panels arranged such that each row is an age range, with the left-hand (right-hand) panel showing the RDD for the school district start (end) date. We simply line up each school district such that the week the school district starts (ends) is at $t = 0$ and we sum reported crimes for each relative calendar week. Patterns for all incidents for 19–24 and 25–30-year-olds are similar to

[19]NIBRS data categorize arrests into Group A and Group B, with Group A arrests including crimes such as drug violations, simple assault, shoplifting, aggravated assault, vandalism, burglary, intimidation, and weapon law violations, and Group B arrests including crimes such as driving under the influence, disorderly conduct, drunkenness, trespassing, and liquor law violations; more than half of Group B arrests are "all other offenses," which can include attempted crimes not completed. Panels A, C, and E of Appendix Figure A.2 shows (Group A) arrests by arrest type, including the "summoned/cited (not taken into custody)" category that we do not include in our main arrest results. Each category exhibits a summer dip, though of differing magnitudes. Panels B, D, and F show similar results for Group B arrests.

[20]Here we note, though, that we find diverging results using a much smaller dataset, the NLSY97 (U.S. Bureau of Labor Statistics, 2022). In Appendix Figure A.4, we aggregate NLSY97 arrests to the monthly level (without weighting), and find that for 10–17 year olds, self-reported arrests are higher in the summer than the in most other months. There is a large spike in June, but also in January, which is puzzling. We additionally considered the month of the most-recent arrest in the older NLSY79 (U.S. Bureau of Labor Statistics, 2023). We found a summer spike, but also unexpected results, such as the number of arrests being roughly only half as large in February as in January, and a large spike in November compared to October. Overall, we place more weight on the evidence from NIBRS, which is an administrative dataset that covers roughly half of the U.S. population, than on the NLSY97 (which includes only about 1,500 10–17 year olds who report being arrested) or the NLSY79 (which includes less than 700 individuals whose most-recent arrest occurred when they were 10–17 years old). The National Crime Victimization Survey, which we discuss below, is also broadly consistent with the findings from NIBRS.

each other (Panels C through F), while the pattern for 10–17 year-olds looks dramatically different. We observe a large increase (roughly 40-50%) in reported crimes involving 10–17 year-olds beginning around the school district start dates (Panel A).[21] There is a large increase in crime about 11–12 weeks before the discontinuity, suggesting that the end of school is also an important driver of changes in reported crime involving children. This is borne out in Panel B, in which we observe similar large (roughly 20-30% lower) drop around the school district end dates (Panel B).[22][23]

In Figure 3, we show similar tables as in Figure 2, but focus on arrests. Again, the older age groups (Panels C–F) show no discontinuity. For 10–17-year-olds, in contrast, we do see a visible discontinuity, particularly around the school end date (Panel B).[24] The discontinuity is present, though less sharp, around the school start date (Panel A).[25] The percentage changes are smaller compared to all incidents in Figure 2.[26] Taken together, the discontinuities around both the beginning and end of school in Figure 2 and 3 provide compelling evidence that school causes an increase in reported crime and arrests.[27]

[21] We present corresponding estimates in Panel A of Appendix Table B.1, which uses a uniform kernel, and in Panel A of Appendix Table B.2, which uses a triangular kernel; we use rdrobust for estimation (Calonico et al., 2017). Each observation is a week. The optimal bandwidth tends to be quite small, such as 3 weeks; we report several different bandwidths. We observe statistically significant coefficients for bandwidths of 6, 8, and 10 weeks. By 10 weeks, there is a 51% increase (uniform kernel) and 42% increase (triangular kernel).

[22] Appendix Tables B.1 and B.2 (Panel B) present the estimates.

[23] See Appendix Figure A.5 for RDDs on other school breaks such as fall break and spring break.

[24] Appendix Figures B.3 and B.4 (Panel B) present the estimates.

[25] Appendix Figures B.3 and B.4 (Panel A) present the estimates. Note that we only detect statistically significant coefficients using the uniform kernel.

[26] In Appendix Figure A.6, we split the arrest results by arrest type for both the school start and school end cutoffs. We find larger jumps for on-view arrests than for arrests from warrants. We also show results for summoned/cited incidents, which we do not consider in our main results, and find much cleaner and larger discontinuities.

[27] In Appendix Figure A.7, we split the 10–17-year-old results by how the incident was reported. The right-hand side columns include only reported crime (arrests) where the date is from the date or report, while the left-hand side columns exclude these. We expect the incidents that rely on date of report, which are an order of magnitude less common, to suffer much more measurement error. Indeed, the right-hand side plots are much noisier. The running variable is, we believe, more often mismeasured, with no visible discontinuity for the arrest results.

3.2 Is This Just Reporting Bias?

In Figure 4, we use the National Crime Victimization Survey (aggregated from 1992–2019) to show that the previous results (based on NIBRS data) are not solely explained by changes in reporting patterns caused by increased scrutiny of crimes committed during the school year. The NCVS microdata contains information from a survey of U.S. households, identifying the members of each household who were victimized by crimes, and the month when that victimization occurred. Between August and September, when most schools begin their school year, we see a large spike in reports of victimization among 10–17-year-olds. And we see a similar drop in reported victimization rates as we move from May to June—lining up with the end of school. We see the opposite summer pattern in criminal victimization for 19–24 and 25–30-year-olds (in Panels B and C): victimization rates spike for these age groups in the summer. This is consistent with the literature on the seasonality of crime. Results are mirrored for reported offender age in the NCVS, based on the recollection of the victims (Panels B, D, and F).

3.3 Heterogeneity

We now return to the NIBRS dataset to better understand the above results by exploring various dimensions of heterogeneity. Below, we explore heterogeneity in results when split by age of offender; hour of the day; whether the reported crime was committed in or outside of school; weekends/weekdays; age of offender/age of victim combinations; type of offense; relationship of offender to victim; reported sex of offender; and reported race of offender. We present all results in the raw data, without focusing on our smaller RDD subsample. Also, for these heterogeneity analyses we focus on all reported criminal incidents (as opposed to only arrests). Unless otherwise specified, we consider offenders aged 10–17.

We begin by looking at heterogeneity by offender age in Appendix Figure A.8, where each panel represents an age, from 10 to 18. We find the same pattern for each age from 10–17, but the pattern disappears—or becomes highly muted—at age 18, when many children are no longer enrolled in high school. Focusing on the trough-to-peak jump around the beginning of school for each age, the percent increase in reported crimes is larger for the youngest ages and smaller for the older ages. There is a more than doubling of reported crimes involving 11–12-year-old offenders, but only about a 20% increase in the reported offenses committed by offenders who are 17 years old.

We next consider the hour of the day the incident occurred. Figure 5 splits the data into the 24 hours of the day. The number of incidents between 1am and 6am peaks in the summer, consistent with general patterns of reported criminal behavior involving adults, but we see the pattern of reported crimes involving 10–17 year olds flip starting at 7am, at which point reported crimes dip in the summer and peaks when the school year begins and ends. The summer dip continues until roughly 8pm, at which point most students are home from school, and the pattern then reverses, returning to the standard summer peak in reported crimes we see for crimes involving adults.[28] This is consistent with the increase in reported crime happening during and immediately after the typical school day, when students are returning home. At first glance, it may seem surprising that the pattern persists several hours after many schools finish. But remember that students may still be involved in activities at school after the school day ends, such as after-school sports, and they may also experience less supervision if their parents are still at work for several hours after the school day ends. In addition, the percent increase in reported crime rates from trough to peak during the calendar year is much larger at, say, 3–3:59pm than it is from 6–6:59pm.

In Figure 6, we recreate the plots of Figure 1, but split the sample by whether the reported crime occurred out-of-school (Panels A, C, and E) or in-school (Panels B, D, and

[28]Note that the number of reported crimes from 12-12:59am appear to suffer from reporting bias. It seems that if the time of day was unknown, it may have been assigned to this time range.

F).[29] The first thing to note is that reported crimes committed in school are orders of magnitude more common for 10–17-year-olds, who are typically enrolled in school, and reported crimes committed in school are much less common for 19–24 and 25–30-year-olds, who are much less likely to be enrolled. The results are striking: all age groups have a large summer decrease in reported crime committed at school and the typical (for adults) summer-peaking seasonal pattern for reported crimes not committed at school. These results show that crimes committed at school explain virtually all of the summer drop (and fall increase) in reported crime for children.[30] While it is not surprising to see the summer dip for crimes committed at school for children (Panel B), the pattern for crimes not at school for children (Panel A) is perhaps surprising. In Appendix Figure A.10, we decompose this graph by the time of day that the incident happened; in Panel A we restrict to what we consider as school hours (between 7am and 7:59pm), and in Panel B we restrict to non-school hours.[31] We find offsetting effects: there is a summer peak in Panel A, but a summer dip in Panel B. In Panels C and D, we repeat this exercise for crimes committed in school. In Panel D, we unsurprisingly find that crimes spike during school hours during the school year (with very few crimes committed during the summer). In Panel C, we find that crimes committed at school, but not during school hours, exhibit a large spike during the beginning of the school year; such crimes are uncommon.

In Appendix Figure A.11, we consider reported crimes occurring on weekdays (Panels A, C, and E) and weekends (Panels B, D, and F) for different age groups. The typical summer-peaking seasonal pattern occurs for older age groups on both weekdays and weekends. In contrast, we find the summer dip for 10–17-year-olds across both types of

[29]We define school as being 'school - elementary/secondary' and 'school/college', but not 'school - college/university.'

[30]In Appendix Figure A.9, we present similar plots using the NCVS victimization data. In Panel A, we observe an increase in the summer, which is somewhat different than what we observe in Panel A of Figure 6. In Panel B, we find a large summer drop for incidents in school, which we also find in Panel B of Figure 6.

[31]We omit midnight to 12:59am because we believe unknown times of reported crimes are often coded in this category, and we also omit 6am to 6:59am because this is an intermediate period between school beginning.

days. However, the percent increase in reported crime rates from summer to the beginning of the fall is much larger on weekdays (more than 50%) versus weekends (roughly 15%).

We proceed to consider the interaction of offender and victim age in Appendix Figure A.12. In particular, we replicate Figure 1, but split the sample by age of offender and age of victim. We focus on two age ranges: 10–17-years-olds; and offenders or victims aged 19 and older. We observe the typical summer-peaking seasonal pattern for reported crimes involving adult offenders and adult victims (Panel D), and a strong summer trough for child offenders and child victims (Panel A). There is a much smaller summer trough for reported crimes involving child offenders and adult victims (Panel B) and we see even less evidence of school-related seasonality for reported crimes involving adult offenders and child victims (Panel C), although this plot is noisier.

We further probe offender and victim age by presenting offender and victim age pairs in Figure 7. For each pair, we calculate a rough ratio of the early fall peak in crime vs. the summer trough in crime by dividing the number of crimes across weeks 37–40 by the number of crimes across weeks 27–30.[32] The darker shades represent higher ratios. Overall, the ratios are highest for same-age offender-victim pairs and for younger offenders and victims. The ratios tend to decrease with with age difference. With some exceptions, we observe higher ratios when the age difference is two years or less. It is likely that our results are at least in part explained by students who are in the same grade as each other and interacting regularly. We expect that results would be strongest for students who are in the same physical school as each other, in accordance with prior evidence on the importance of peers and network effects in reported criminal activity (for example, Billings et al. (2019). We see some evidence of this, with drop offs between ages 13 and 14 (for many children, roughly middle vs. high school), and between 10 and 11 (for many children, roughly elementary vs. middle school) and between 17 and 18 (high school vs.

[32]We include the full plots from which the heatmap is derived for many of the the age pairs in Appendix Figure A.13 (offenders aged 12–14) and Appendix Figure A.14 (offenders aged 15–17).

not).

Next, in Figure 8, we split by the number of offenders involved in the incident. Our results are driven entirely by incidents involving either one (Panel A) or two offenders (Panel B), with a larger relative summer drop in the former than the latter. The pattern completely flips for three or more offenders (Panel C), which exhibits a summer peak.

In Figures 9, we show that these patterns are explained by several offense categories. After classifying UCR codes into ten groups of crimes, we show that the beginning of the school year causes a sharp uptick in reported drug crimes and simple assaults, as well as reported non-violent crimes like intimidation and weapons law violations. We see either a weaker pattern (meaning a small percentage difference between trough and peak) or do not see this pattern at all for killings; theft; sexual assault; and property damage or vandalism. For comparison, in Appendix Figure A.17, we show NCVS results for 10-17 year old victims split by our classification of victimization codes into type of offense. We find roughly similar results: we observe large summer drops for simple assault, verbal threats, and assaults involving no weapon and no injury. There is also a drop for theft (not present in NIBRS), but small or no drops for other violent and sexual assaults.

We next split the sample by the relationship of the offender to the victim in Figure 10. Reported crimes in which the offender is an acquaintance to the victim are both most common and have the strongest summer trough pattern in terms of percent increases in reported crimes. This pattern also appears to varying degrees for the other groups, being strongest among friends, romantic relationships, and the uncommon 'the victim is the offender' category).[33] It is much much weaker among family members and strangers.

In Appendix Figures A.15 and A.16, we examine these trends separately for males and females, as well as by the race categories law enforcement agencies classify offenders and victims into.[34] We see the same general patterns across sex-based and race-based

[33]Victim is offender can occur in instances such as brawls, where an individual both attacks and is attacked.

[34]The race categories we use are White, Black, and Other Race, into which we group American Indian, Alaskan Native, Asian, Native Hawaiian, and Pacific Islander.

groups. Surprisingly, for offenders and victims among all categories—males, females, and each racial classification—we find roughly the same proportional increase in reported crime, approximately 40–50%, surrounding the start of the school year. This indicates that the school environment has a proportional **percent** increase in reported crimes involving each group. As we discuss in more detail later, these proportional percent increases in reported crimes for different demographic groups are percent increases from different levels, meaning that the school environment causes sharp increases in the Black-white gap and male-female gap in reported crimes and arrests.

3.4 Case Study: School-based Police Agencies

To explore the mechanisms underlying these effects, we now consider the phenomenon of school districts that have their own law enforcement agency. As a case study, we consider the city of Austin, Texas and the Austin Independent School District (ISD), which have distinct law enforcement agencies that operate separately and report separately to NIBRS.[35] In Figure 11 Panel A, we plot the reported crimes by week for 10–17-year-old offenders in the Austin ISD, which provides law enforcement support to the public school system. We see a sharp drop in reported crime during the summer weeks, with reported crimes during these weeks being zero or close to it. Around calendar week 35, there is a sharp uptick in reported crime. This is in contrast to the Austin police department (Panel B), which has a relatively uniform rate of reported crime involving 10–17-year-old offenders over the year (though there is a small dip at the end of the summer). In Panel C, we show reported crime for a broader sample of police agencies that include the word "school," "schools," "isd", or "i.s.d." in their title. We see the same pattern as in Panel A. In Panel D, we present a similar graph for the police agencies in the same counties as the school-based police agencies in Panel C; we further restrict to police agencies that have the same first word in their name as do school districts within the county; while this is

[35]We show data beginning in March 2019, when Austin ISD first started reporting its data.

not perfect, we do this because it is common for there to be multiple police agencies in the same county, some of which would not correspond to the school district. Keeping in mind that there are few observations, we find little or no discontinuity. For comparison, in Panel E, we plot weekly reported crime for colleges and university police agencies (that include the word "college" or "university" in their title; age is unrestricted). We see a similar pattern, though the baseline level of reported crime in summer months is much greater than for school police agencies.[36]

3.5 Distance to Nearest School - New York City

We now turn our attention to New York City, which records the latitude and longitude of its reported crimes in incident-level microdata. We associate each reported crime for offenders 17-years-old and younger to the nearest school in New York City, and we present weekly reported crime by distance. In Figure 12, we find a strong summer dip in reported crime—a more than halving—for reported crimes within 0.05 miles of the nearest school (Panel A). We observe a similar, but greatly dampened pattern for reported crimes between 0.05 and 0.1 miles and between 0.1 and 0.25 miles from the nearest school (Panels B and C). The effect, if any, is small and noisy further than 0.25 miles from the nearest school.

In Appendix Figure A.18, we show that reported crimes for offenders 17 and younger occurring within 0.1 miles from the nearest school show a stronger pattern for public schools (Panel A) than for private schools (Panel B). The pattern is similar in terms of percent change from trough to peak across types of school when we split our sample into schools with many reported crimes (Panel C) and schools with few reported crimes (Panel D). In other words, there is no clear pattern where the schools that have the highest levels of reported crime also see the largest evidence of seasonality in reported crime nearby to the school.

[36] Colleges and universities also tend to have more students enrolled during the summer.

3.6 Excess Crime Calculation

In the previous section, we presented evidence that the school environment causes an increase in arrests and reported crimes among 10–17-year-olds; But we have done little to quantify the magnitude of these effects. To approach this question systematically, we return to the 2017–2019 NIBRS data. Within this data, we first calculate an 'expected' number of reported crimes and arrests based on seasonality and demographics. Our procedure for this is as follows. We construct daily counts of reported crimes and arrests for cells defined by the interaction of calendar day (January 1st, 2017 through December 31st, 2019), hour-of-the-day (1am to 2am,... through 11pm to midnight)[37], and age of offender (ages range in this estimation sample from 10 through 40). As an example, one of our cells is the number of reported crimes committed on January 1st, 2019, between the hours of 1am and 2am, by 10-year-olds. We then estimate the following regression at the cell-level, after **subsetting to** cells reported outside of the school calendar and school hours:[38]

$$Y_{a,t,h} = \alpha_{a,t} + \delta_{a,dow(t),h} + \epsilon_{a,t,h},$$

where $Y_{a,t,h}$ is a count of one of two outcomes—reported crimes or arrests in NIBRS—measured for offenders of age a on calendar day t during hour-of-the-day h. $\alpha_{a,t}$ represents age -by- calendar day fixed effects, and $\delta_{a,dow(t),h}$ represents day-of-the-week -by-hour-of-the-day -by- age fixed effects. This regression produces a set of seasonal and demographic-based adjustments that we can then use to residualize our counts of arrests and reported crimes at the cell-level throughout the year.

We then take the residualized outcome $\Delta_{a,t,h} = Y_{a,t,h} - \hat{Y}_{a,t,h}$ and plot the average residual count of reported crimes or arrests by calendar day and age.[39] Put another way,

[37]We drop observations from midnight to 1am due to what appears to be large measurement error due to assigning crimes with an unknown time to this time. We also drop observations that are missing at least one of the age, date, and hour variables. These dropped observations are excluded from all calculations below.

[38]This includes cells from July and cells in non-July months covering the hours of 8pm to 6am.

[39]It is important to remember that our $\hat{Y}_{a,t,h}$ is constructed using the fixed effects from the above equation,

we are residualizing out both daily seasonality of reported crime, which we allow to vary by age and which we estimate using evening and early morning hours (estimated using $\alpha_{a,t}$), and hourly seasonality of crime, which we allow to vary by day-of-the-week and age, and which we estimate using reported variation in criminal activity throughout the day in the month of July.

In Figure 13 we plot these residuals, summed across all cells within a calendar date ($\bar{\Delta}_{a,t,\cdot}$). The top panel shows the residuals for weekdays, and the bottom panel shows residuals for weekends. In both panels, for each age group we peg the daily residual to be zero in July. Here we see a clear pattern. For non-school-age offenders, the average daily value of these residuals is approximately zero. In other words, the patterns of seasonality for non-school-age offenders during school hours of the year is well-predicted by the patterns of seasonality that we estimate outside of the school year. But, for school-aged children (10–17-year-olds), we see a different pattern. The patterns of seasonality outside the school day do not do a good job of explaining the amount of reported crime during the school day. Instead, during the school year we see elevated counts of reported crime relative to what we would predict based purely on our seasonality estimates. As a placebo test, we can perform the same exercise for weekends (bottom panel), plotting the average residual for each weekend. Here, we no longer see as clear of a pattern. While there are somewhat higher average residuals for 10–17-year-olds, perhaps reflecting the increased interactions of students with peers on weekends during the school year (see Panel B of Figure A.11, there is much more overlap with the older ages than in Panel A.

To quantify the magnitude of these 'excess' reported crimes for 10–17-year-olds, on the average day of the year (including July and weekends), we see 297 excess reported crimes in our 2017–2019 NIBRS data, relative to what we would expect from non-school-hour patterns of seasonality. As a comparison group, we can look at 25–30-year-olds[40],

estimated only on cells from July and from outside-of-July that are also outside of school hours.

[40]25–30 year olds are less likely to be in college than are 19–24 year olds; we use 19–24 year olds as the comparison group in a robustness check: Appendix Table B.5 shows qualitatively similar results, with slightly smaller estimates.

who are on average reported as having committed 52 additional crimes on a given day, relative to what we might expect from non-school-hour seasonality patterns. Because there are many more crimes reported involving the 25–30 group, we re-scale the comparison number, 52, by multiplying by the total number of crimes by 10–17 year-olds divided by the total number of crimes by 25–30 year-olds; this number becomes 26.[41] So, to perform a simple back-of-the-envelope calculation, we estimate that the school environment is responsible for $297 - 26 = 271$ excess reported crimes per day, or an excess $271 \times 365(days) \times 3(years) = 296,745$ reported crimes.[42] This is 32% of all 933,676 crimes reported to NIBRS involving 10–17-year-old offenders in 2017–2019. Or, said another way, this calculation implies that the school environment increases reported crimes with 10–17-year-old offenders by $\frac{296,745}{933,676-296,745} = 47\%$ (Table 13).

We can perform the same exercise for arrests (see Figure 14). We find smaller but similar effects. On the average day of the year, we see 71 excess arrests for 10–17 year olds in our sample, relative to what we would expect from non-school-hour patterns of seasonality. As a comparison group, we can look at 25–30-year-olds, who are on average reported as having 15 additional arrests on a given day, relative to what we might expect from non-school-hour seasonality patterns; after scaling the 25–30 year-old group to have the same number of arrests as the 10–17 year-old group, our comparison number is 5. This implies that in 2017–2019, the school environment is responsible for $71 - 5 = 66$ excess arrests per day, or an excess $66 \times 365 \times 3 = 72,270$ arrests each year. This is 29% of the 247,776 arrests reported to NIBRS involving 10–17-year-old offenders in 2017–2019. In other words, we argue that the school environment increases arrests among 10–17-year-olds by $\frac{72,270}{247,776-72,270} = 41\%$.

We now separately consider how this excess crime measure varies by the reported

[41]In other words, we adjust the raw number of crimes involving 25–30 year-olds down by roughly a factor of two to account for the fact that there are around twice as many reported crimes involving that age group in our sample.

[42]This number is approximately correct, but is obtained using the rounded numbers in the text used for clarity.

race and sex of offenders. Here, we find remarkably similar percent increases in crime for each group of offenders. Using the same method reported previously, in Appendix Figure A.19, we show our daily residual (excess) reported crime on weekdays for male offenders (Panel A), female offenders (panel B), Black offenders (panel C), and White offenders (panel D). Here we see patterns consistent with the results from the prior two paragraphs. In all cases, we see excess reported crime for 10–17-year-olds relative to what we would expect to see from seasonal patterns in crime outside of school hours. In Appendix Figure A.20, we plot the same residuals but on weekends, and while the 10–17 residuals are elevated, this elevation is relatively small in magnitude.

To quantitatively estimate the number of excess reported crimes due to the school environment, we can now perform the same back-of-the-envelope calculation described above, but separately for male, female, Black, and White offenders. We present results for all incidents in Panel A of Table 1 and for arrests in Panel B of this same table. We find that the school environment causes 32% of all crime involving 10–17-year-olds, among White, Black, male, and female offenders.[43] This means that for each of those demographic groups, the school environment is causing a 46-47% increase in annual reported crimes. The analogous numbers for arrests imply that the school environment causes 28–30% of all arrests involving 10–17 year-olds, increasing arrest rates by 39–42% for each group.[44] This may seem somewhat equitable on its face. Whatever is happening during the school day—whether it is an increase in reported crimes because school creates additional opportunities for children to interact with peers or additional factors—it is leading to about the same percentage increase across groups. *But* this 47% increase in reported crimes (or 41% for arrests) are increases relative to different baseline levels of reported criminal activity for males, females, Black, and white (reported) offenders; so, the school environment significantly magnifies any preexisting inequities in the amount of race-based or sex-based reported crime or arrest rates.

[43] We also find similar numbers for offenders of other races.
[44] We find higher numbers for arrestees of other races.

To see why this is the case, consider a world with two groups, A and B. In Group A, there are 100 reported crimes with child offenders per 1,000 children annually. In Group B, there are 10 reported crimes with child offenders per 1,000 children annually. This means that there is an A-B gap of 90 reported crimes per 1,000 children annually. If an intervention increases the number of reported crimes per 1,000 children by 47% for each of Group A and Group B, then in the new state of the world, Group A will report 147 crimes per 1,000 children and Group B will report 14.7 crimes per 1,000 children. So the new A-B gap in reported crime rates is 147-14.7 = 132.3 reported crimes per 1,000 children. In other words, the gap has increased by 47%.[45]

The reason we do not calculate the change in the Black-White and male-female reported per-capita crime rates directly here is because the denominator is not easy to calculate. NIBRS data covers around half of the U.S. population, but students in schools do not need to attend schools within the same police agency as their place of residence, although they often do. But, applying the results described above, we can say that no matter what the baseline levels of reported crime and arrest rates are among male, female, Black, and white children, the school environment increased both the male-female and Black-white gaps in reported crime rates by around 47%. And the school environment increases the male-female and Black-white gaps in arrest rate by around 41%.

3.7 Predictors of Excess Crime

In our final analysis, we explore whether our national measure of excess reported crime caused by time spent in school varies as a function of characteristics of schools and places. We first calculate county-level measures of excess crime . We then create a series of county-level indicators and potential predictors that one would theoretically expect to affect excess crime. These potential predictors include the poverty rate, the

[45]This is seen more generally in a world where group A has X reported crimes per 1,000 children and group B has Y reported crimes per 1,000 children. If an intervention increases reported crime rates by Z%, then the A-B gap in reported crimes increases from $X - Y$ to $(1 + Z)X - (1 + Z)Y = (1 + Z)(X - Y)$.

unemployment rate, median household income, population (in levels), the percentage of adults with less than a high school education, rurality, spending per student, the number of students per grade among 5–12 graders, the percentage of grade 5–12 students who are white or Asian, the configuration of a county's schools (the fraction of high schoolers, or 9–12th graders, also in school with middle-schoolers, or 6–8 graders), and the presence of a school resource office or security guard. We regress the county-level measure of excess crime on the set of predictors described above for the 1,824 counties we consider.[46]

As we can see in Table 2, which weights by county population, none of the predictors survives state-level fixed effects and clustering at the state level (column 4). This indicates that our main excess crime results hold across places of different types. In Tables B.6–B.8, we present similar tables, but don't weight, weight by total grade 5–12 school enrollment, and weight by the number of reported crimes among 10–40 year olds (after sample restrictions). With only three exceptions—two of which are only marginally significant—the predictors are still uncorrelated with local excess crime measures in these three tables. In other words, we find no consistently strong predictors of local variation in excess crime.

4 Conclusion

In this paper, we document a sharp increase in arrests and reported criminal activity involving children as the school year starts, and a sharp decline in both measures as the school year ends. We use a regression discontinuity design to show that this pattern is the result of a causal relationship: as students begin to attend school in the end of summer and early fall, crime rates among children spike because of arrests and reported crimes committed during school hours. We use data from a victimization survey to show that these patterns are not explained by changes in crime reporting patterns during the school

[46]See the Data Appendix for details on how we constructed the variables and selected the sample.

year.

Our findings are consistent across demographic groups and counties of different types, emphasizing the universal impact of school on criminal activity involving children. Because of these consistent patterns, we show that the school environment increases reported crime rates among 10–17-year-olds by 47%, and also increases the Black-white gap in reported crimes by roughly the same amount.

While our results challenge the prevailing understanding of the benefits of education in reducing criminal activity among children, we argue that our results are not inconsistent with those findings. Much of the work linking increased education to reductions in crime focuses on policy changes that affect children on the margin of dropping out of school, a sub-population that may be especially positively affected by a focus on school. What our paper shows is a population-level outcome: across all children—those who are compliers for education-based policies and those who are not—school causes an average increase in the likelihood of interacting with the criminal justice system.

Looking forward, we hope to explore the underlying mechanisms driving the observed relationship between the school environment and juvenile criminal activity. Future research may unravel specific aspects of the school environment that drive these effects, such as peer interactions, disciplinary practices, and school policies. And a better understanding of these mechanisms will guide the development of targeted interventions and policy reforms aimed at reducing children's interactions with the criminal justice system both within and outside of school. We also leave as an open question the long-run effects of school-caused engagement with the criminal justice system on adult outcomes: a series of analyses which may highlight additional costs of this aspect of school.

Ultimately, such insights will guide the development of targeted interventions and policy reforms aimed at fostering a safer and more conducive learning environment for children, thereby mitigating the unintended consequences of school-based interactions with the criminal justice system.

References

Akee, R. Q., T. J. Halliday, and S. Kwak (2014). Investigating the effects of furloughing public school teachers on juvenile crime in hawaii. *Economics of Education Review 42*, 1–11.

American Institutes for Research (2020). Crdc and edfacts id crosswalk. https://www.air.org/project/research-evaluation-support-civil-rights-data-collection-crdc.

Anderson, D. M. (2014). In school and out of trouble? the minimum dropout age and juvenile crime. *Review of Economics and Statistics 96*(2), 318–331.

Bacher-Hicks, A., J. Goodman, J. G. Green, and M. Holt (2022). The covid-19 pandemic disrupted both school bullying and cyberbullying. Technical report, National Bureau of Economic Research.

Bell, B., R. Costa, and S. Machin (2022). Why does education reduce crime? *Journal of political economy 130*(3), 732–765.

Billings, S. B., D. J. Deming, and S. L. Ross (2019). Partners in crime. *American Economic Journal: Applied Economics 11*(1), 126–50.

Block, C. R. (1984). *Is crime seasonal?* Illinois Criminal Justice Information Authority Chicago.

Bureau of Justice Statistics (2018). Law enforcement agency identifiers crosswalk, united states, 2012. *Inter-university Consortium for Political and Social Research [distributor]*. https://doi.org/10.3886/ICPSR35158.v2.

Bureau of Justice Statistics. (2021). National crime victimization survey, concatenated file, [united states], 1992-2020. *Inter-university Consortium for Political and Social Research [distributor]*.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal 17*(2), 372–404.

Carbone-Lopez, K. and J. Lauritsen (2013). Seasonal variation in violent victimization: Opportunity and the annual rhythm of the school calendar. *Journal of quantitative criminology 29*(3), 399–422.

Cowan, B. W., T. R. Jones, and J. M. Swigert (2023). Parental and student time use around the academic year. Technical report, National Bureau of Economic Research.

Edunomics Lab at Georgetown University (2021). National education resource database on schools (nerd$), 2018-19. https://edunomicslab.org/.

Fischer, S. and D. Argyle (2018). Juvenile crime and the four-day school week. *Economics of Education Review 64*, 31–39.

Fitzpatrick, M. D., C. Benson, and S. R. Bondurant (2020). Beyond reading, writing, and arithmetic: The role of teachers and schools in reporting child maltreatment. Technical report, National Bureau of Economic Research.

Hansen, B. and M. Lang (2011). Back to school blues: Seasonality of youth suicide and the academic calendar. *Economics of Education Review 30*(5), 850–861.

Hansen, B., J. J. Sabia, and J. Schaller (2022). In-person schooling and youth suicide: Evidence from school calendars and pandemic school closures. Technical report, National Bureau of Economic Research.

Hjalmarsson, R., H. Holmlund, and M. J. Lindquist (2015). The effect of education on criminal convictions and incarceration: Causal evidence from micro-data. *The Economic Journal 125*(587), 1290–1326.

HUD PDR (2017). Hud-usps zip crosswalk files. ZIP-County, 2nd Quarter 2017, https://www.huduser.gov/apps/public/uspscrosswalk/home.

Jacob, B. A. and L. Lefgren (2003). Are idle hands the devil's workshop? incapacitation, concentration, and juvenile crime. *American economic review 93*(5), 1560–1577.

Kaplan, J. (2021a). Jacob kaplan's concatenated files: National incident-based reporting system (nibrs) data, 1991–2019. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research.[distributor]*.

Kaplan, J. (2021b). *National Incident-Based Reporting System (NIBRS) Data: A Practitioner's Guide*. https://nibrsbook.com/.

Luallen, J. (2006). School's out... forever: A study of juvenile crime, at-risk youths and teacher strikes. *Journal of urban economics 59*(1), 75–103.

Machin, S., O. Marie, and S. Vujić (2011). The crime reducing effect of education. *The Economic Journal 121*(552), 463–484.

McDowall, D., C. Loftin, and M. Pate (2012). Seasonal cycles in crime, and their variability. *Journal of quantitative criminology 28*(3), 389–410.

National Center for Education Statistics (2020). School locations geoassignments 2019–2020. https://nces.ed.gov/programs/edge/Geographic/SchoolLocations.

New York City Police Department (2022). NYPD complaint data historic. https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i.

Quetelet, A. (1842). *Instructions pour l'observation des phénomènes périodiques (in English: A Treatise on Man and the Development of His Faculties)*. English translation 1968. New York: Burt Franklin.

Shulman, E. P., L. D. Steinberg, and A. R. Piquero (2013). The age–crime curve in adolescence and early adulthood is not due to age differences in economic status. *Journal of Youth and Adolescence 42*(6), 848–860.

Urban Institute (2023). Education data portal (version 0.18.0). https://educationdata.urban.org/documentation/, made available under the ODC Attribution License.

U.S. Bureau of Labor Statistics (2022). National longitudinal survey of youth 1997. https://www.nlsinfo.org/content/cohorts/nlsy97.

U.S. Bureau of Labor Statistics (2023). National longitudinal survey of youth 1979. https://www.nlsinfo.org/content/cohorts/nlsy79.

U.S. Department of Education (2018). Civil rights data collection 2017-18. https://ocrdata.ed.gov/resources/downloaddatafile.

USDA Economic Research Service (2019a). Educational attainment for the u.s., states, and counties, 1970-2018. https://web.archive.org/web/20200410144520/https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/.

USDA Economic Research Service (2019b). Population estimates for the u.s., states, and counties, 2010-18. https://web.archive.org/web/20200410144520/https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/.

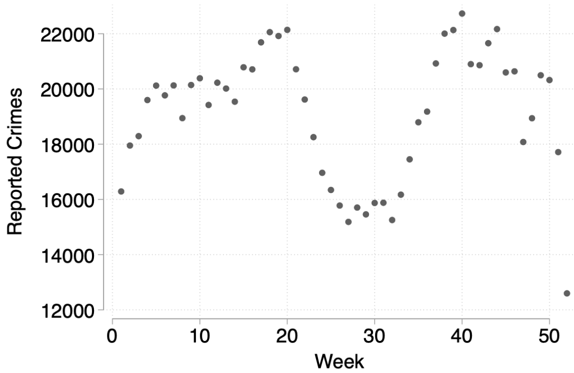
USDA Economic Research Service (2019c). Poverty estimates for the u.s., states, and counties, 2018. https://web.archive.org/web/20200410144520/https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/.

USDA Economic Research Service (2019d). Unemployment and median household income for the u.s., states, and counties, 2007-18. https://web.archive.org/web/20200410144520/https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/.

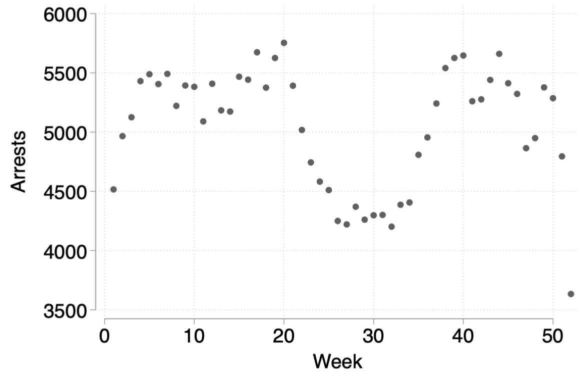
USDA Economic Research Service (2020). 2013 rural-urban continuum codes. https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/.

Wang, H., K. Paulson, S. Pease, S. Watson, H. Comfort, P. Zheng, A. Aravkin, C. Bisignano, R. Barber, T. Alam, and J. Fuller (2022). Estimating excess mortality due to the covid-19 pandemic: a systematic analysis of covid-19-related mortality, 2020–21. *The Lancet 399*, 1513–1536.

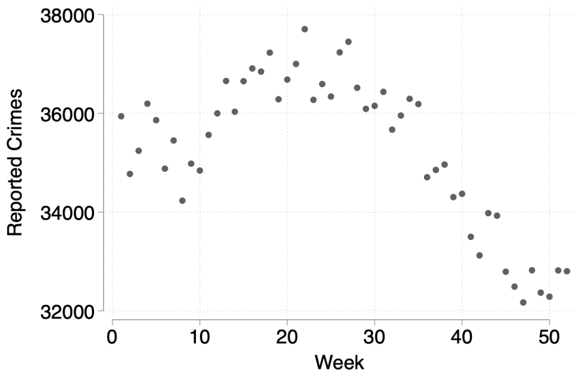
Figure 1: Reported Crimes and Arrests by Week



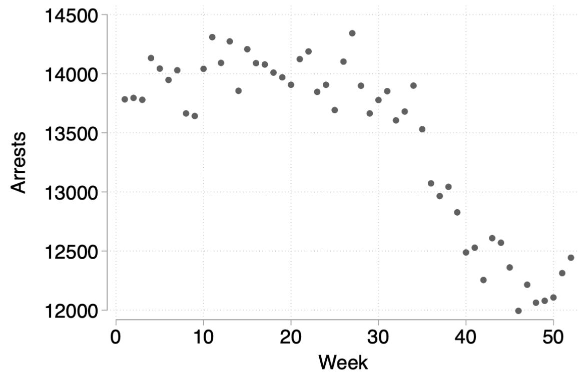
(a) 10-17, All Incidents



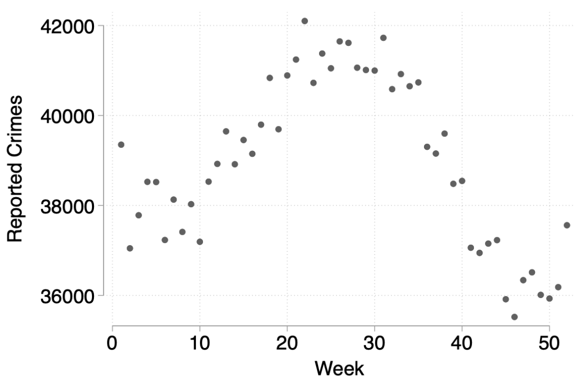
(b) 10-17, Arrests



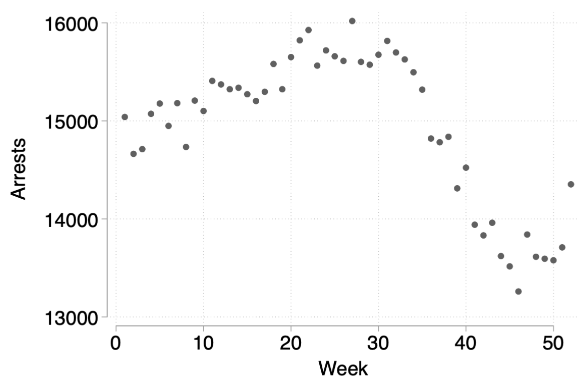
(c) 19-24, All Incidents



(d) 19-24, Arrests



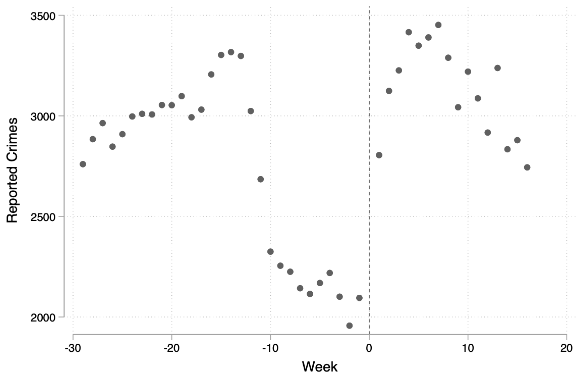
(e) 25-30, All Incidents



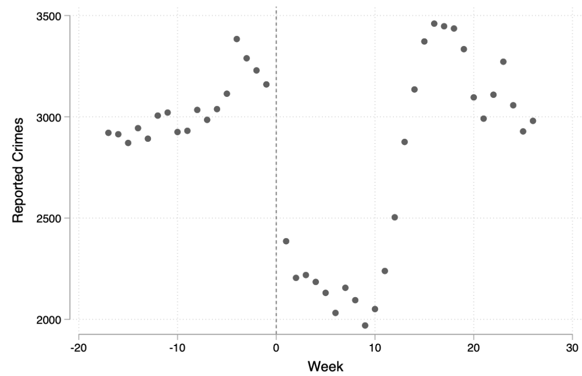
(f) 25-30, Arrests

Notes: These figures show the total number of reported crimes and arrests by week. Panels A, C, and E show the total number of reported crimes, and Panels B, D, and F show the total number of arrests. Panels A and B are limited to the offender being 10-17 years old; Panels C and D (E and F) are limited to the offender being 19-24 (25-30) years old. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

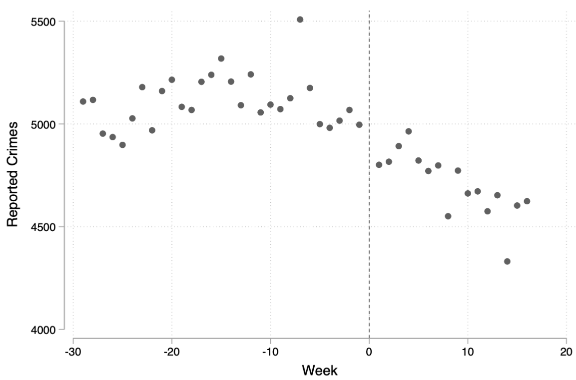
Figure 2: Reported Crimes by Week Relative to School Start Date



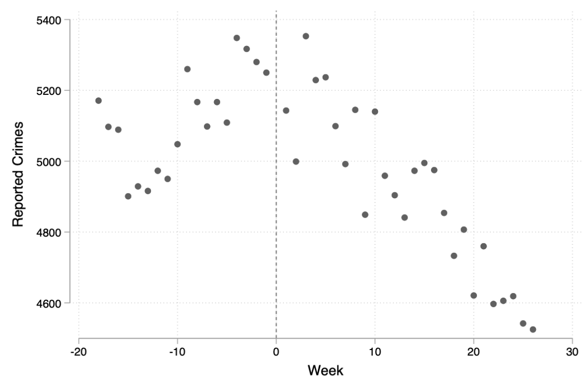
(a) 10-17, School Start



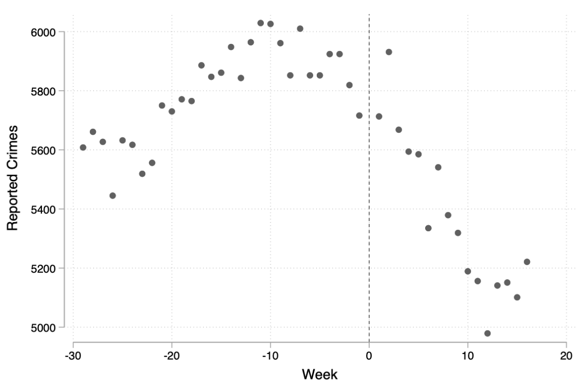
(b) 10-17, School End



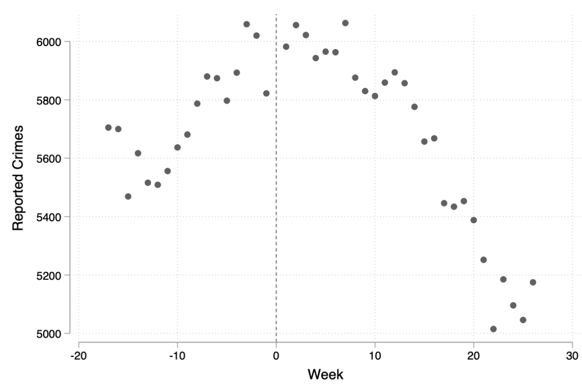
(c) 19-24, School Start



(d) 19-24, School End



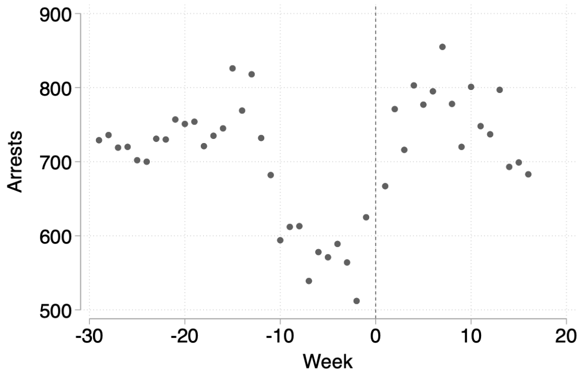
(e) 25-30, School Start



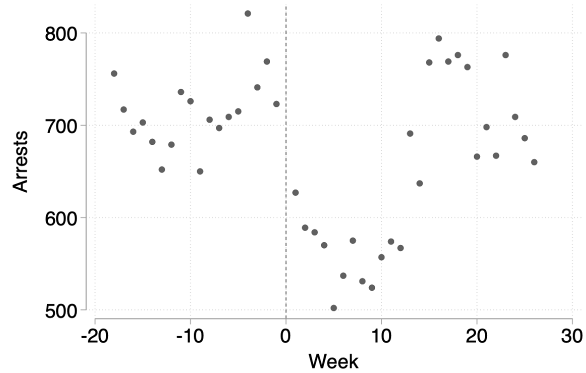
(f) 25-30, School End

Notes: These figures show a regression discontinuity for the number of reported crimes by week relative to school district start date (Panels A, C, and E) and school district end date (Panels B, D, and F). Panels A and B are limited to the offender being 10-17 years old; Panels C and D (E and F) are limited to the offender being 19-24 (25-30) years old. The data source is 2019 NIBRS. Week 0 is the school district starting (ending) date. Week 0 is excluded as this is sometimes a partially treated week. A line is drawn at week 0.

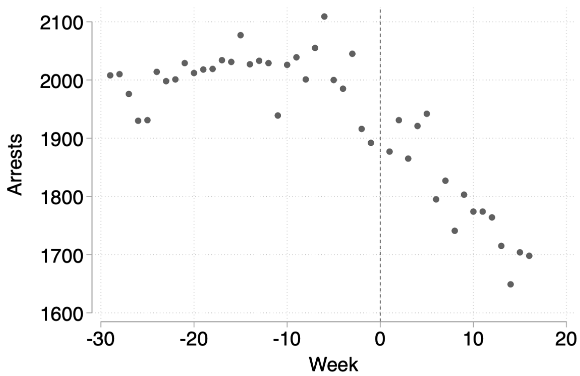
Figure 3: Arrests by Week Relative to School End Date



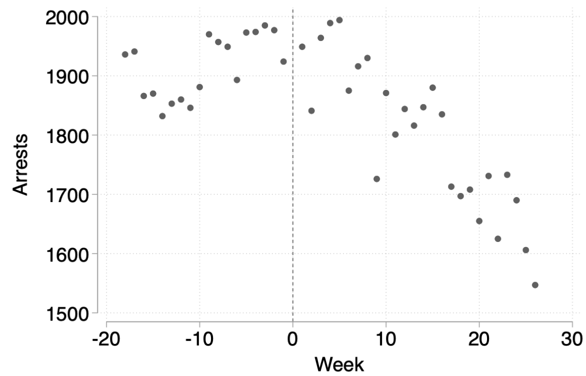
(a) 10-17, School Start



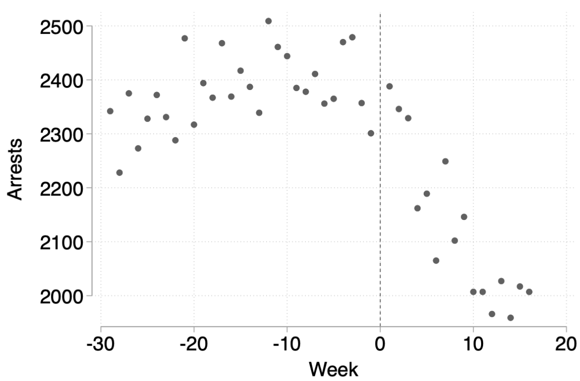
(b) 10-17, School End



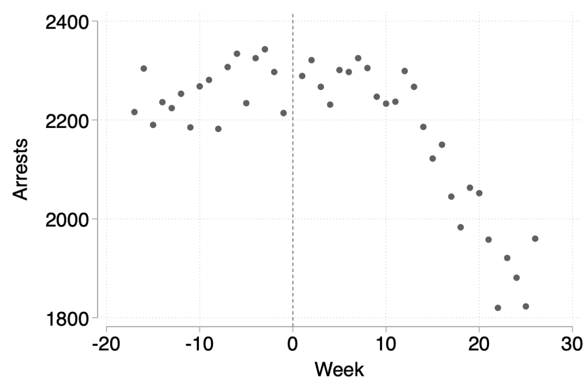
(c) 19-24, School Start



(d) 19-24, School End



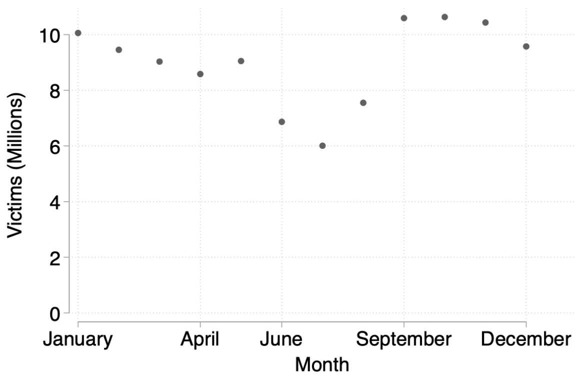
(e) 25-30, School Start



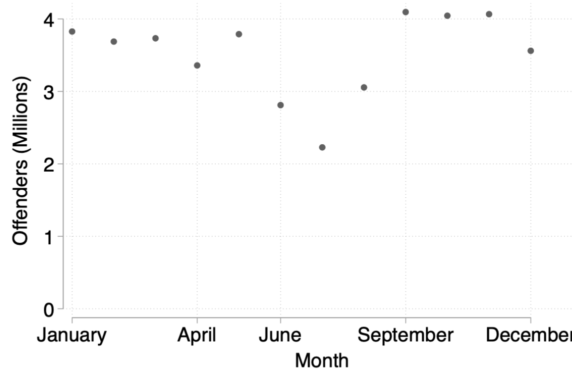
(f) 25-30, School End

Notes: These figures show a regression discontinuity for the number of arrests by week relative to school district start date (Panels A, C, and E) and school district end date (Panels B, D, and F). Panels A and B are limited to the offender being 10-17 years old; Panels C and D (E and F) are limited to the offender being 19-24 (25-30) years old. The data source is 2019 NIBRS. Week 0 is the school district starting (ending) date. Week 0 is excluded as this is sometimes a partially treated week. A line is drawn at week 0.

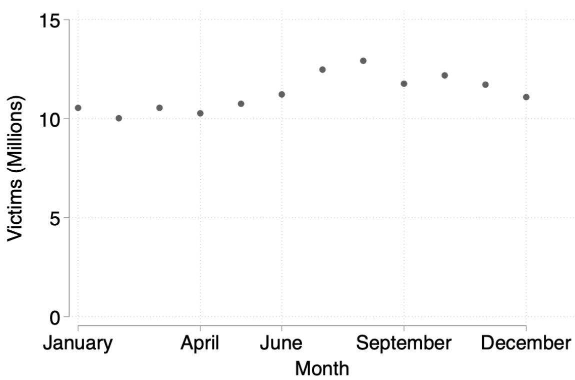
Figure 4: Crimes by Month, National Crime Victimization Survey Data, by Victim/Offender



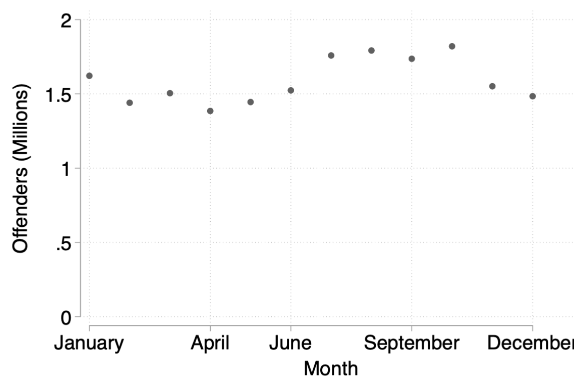
(a) Victim Age 10-17



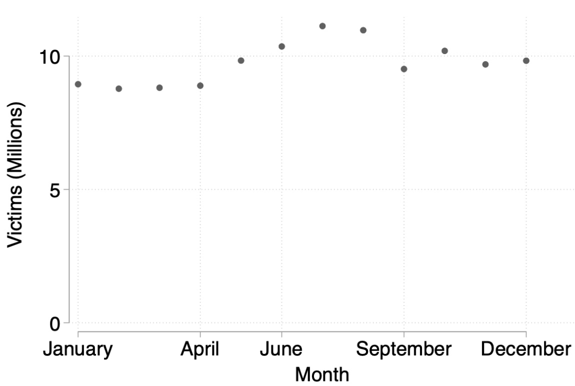
(b) Offender Age Under 18



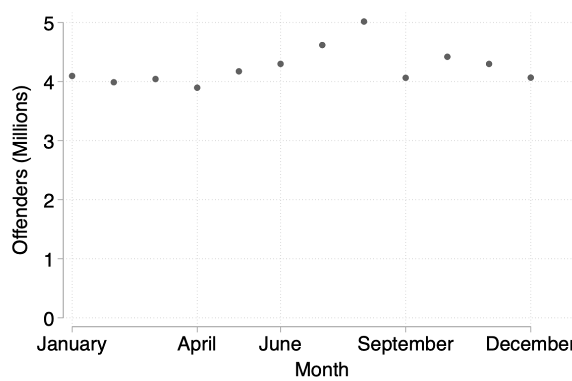
(c) Victim Age 19-24



(d) Offender Age 18-20



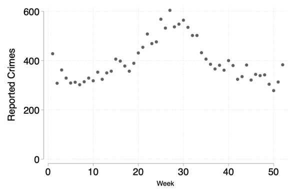
(e) Victim Age 25-30



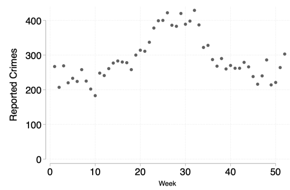
(f) Offender Age 21-29

Notes: These figures show the number of crimes by month using data from the National Crime Victimization Survey from 1992 to 2019 (after weighting using person weights). Panels A, C, and E are for victims, and Panels B, D, and, E are for offenders. Panel A is limited to the victim being 10-17 years old; Panel C (E) is limited to the victim being 19-24 (25-30) years old. Panel B is limited to the offender being 17 and younger; Panels D (E) is limited to the offender being 18-20 (21-29) years old. The offender age bin is often unobserved, leading to few observations for offenders than for victims.

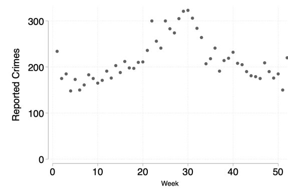
Figure 5: Crimes by Week, Age 10-17, by Hour of the Day



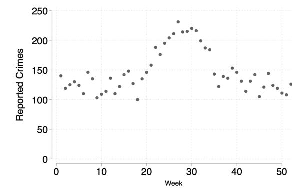
(a) 1-1:59am



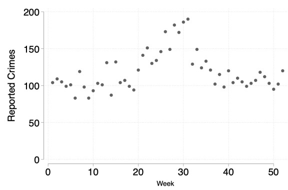
(b) 2-2:59am



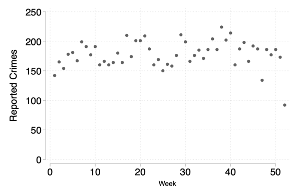
(c) 3-3:59am



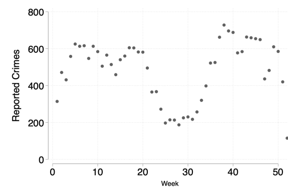
(d) 4-4:59am



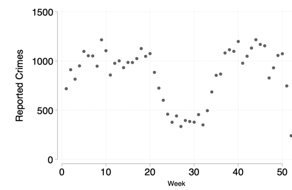
(e) 5-5:59am



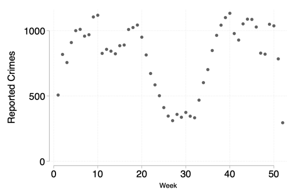
(f) 6-6:59am



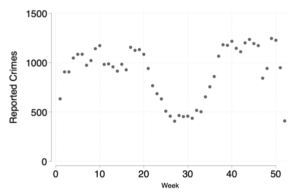
(g) 7-7:59am



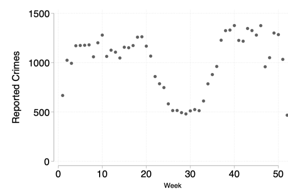
(h) 8-8:59am



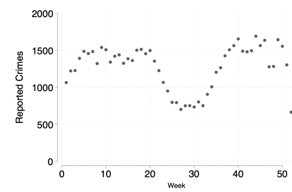
(i) 9-9:59am



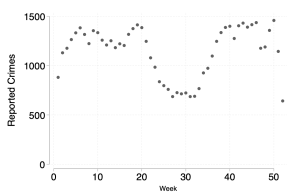
(j) 10-10:59am



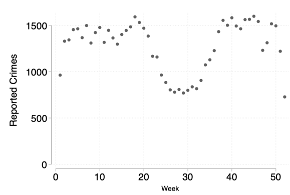
(k) 11-11:59am



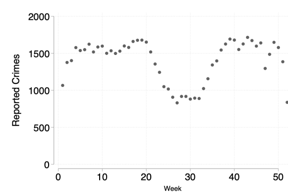
(l) 12-12:59pm



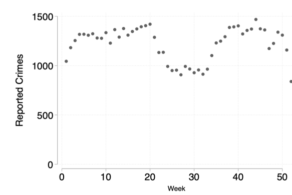
(m) 1-1:59pm



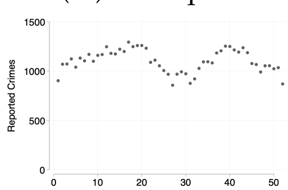
(n) 2-2:59pm



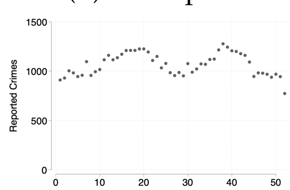
(o) 3-3:59pm



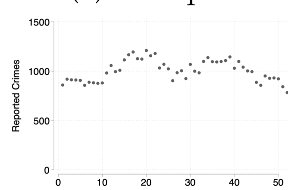
(p) 4-4:59pm



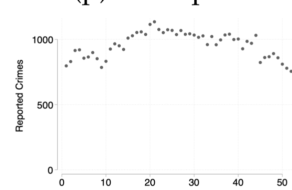
(q) 5-5:59pm



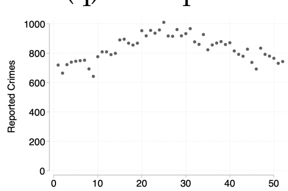
(r) 6-6:59pm



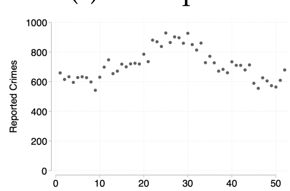
(s) 7-7:59pm



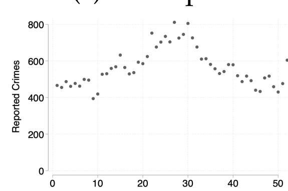
(t) 8-8:59pm



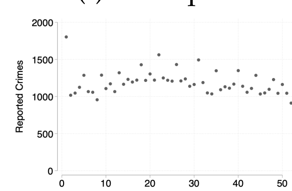
(u) 9-9:59pm



(v) 10-10:59pm



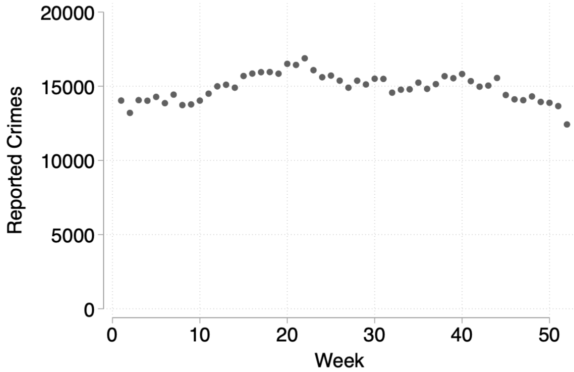
(w) 11-11:59pm



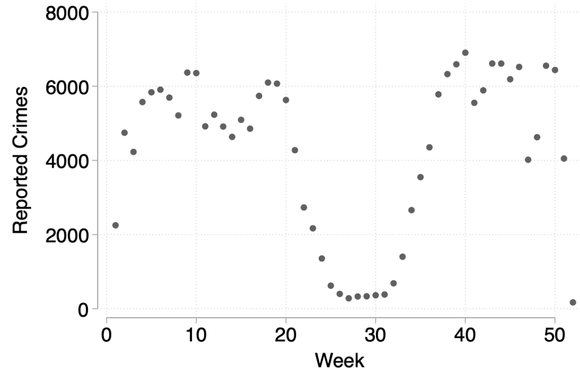
(x) 12-12:59am

Notes: These figures show the total number of reported crimes by week for ages 10-17. Each panel corresponds to a different hour of the day. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

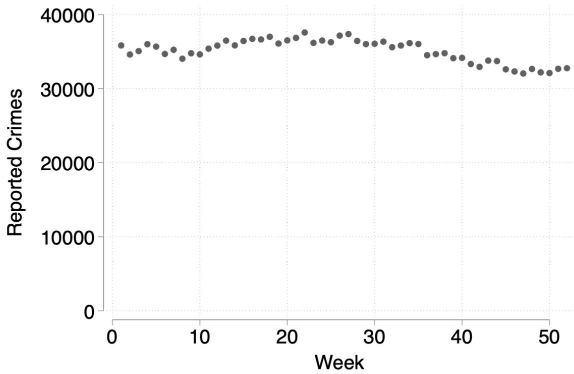
Figure 6: Reported Crimes by Week, by Location



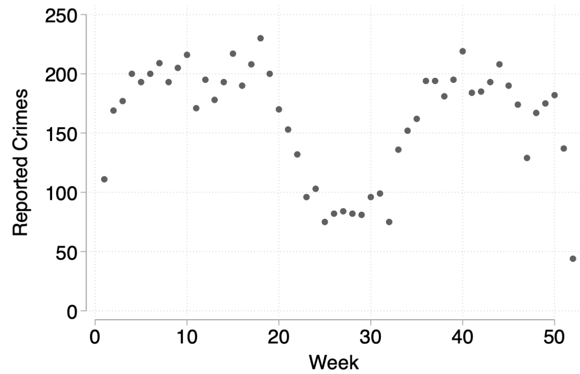
(a) Not in School; 10-17



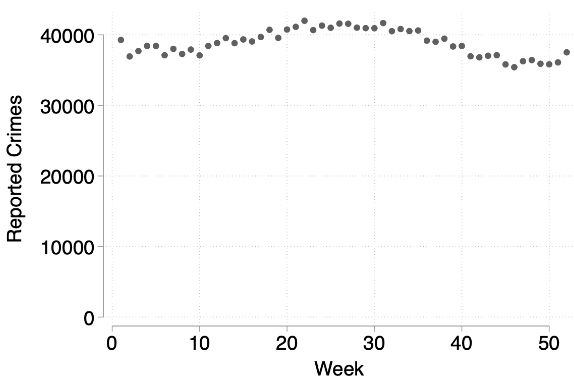
(b) In School; 10-17



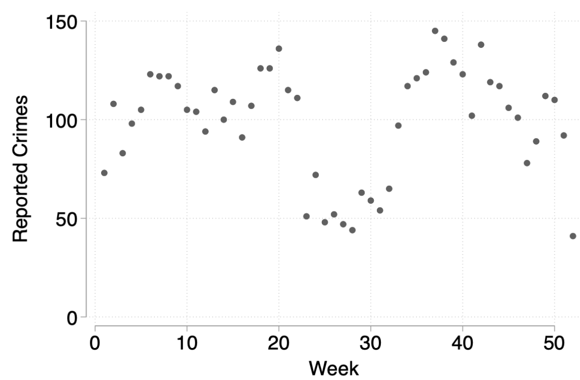
(c) Not in School; 19-24



(d) In School; 19-24



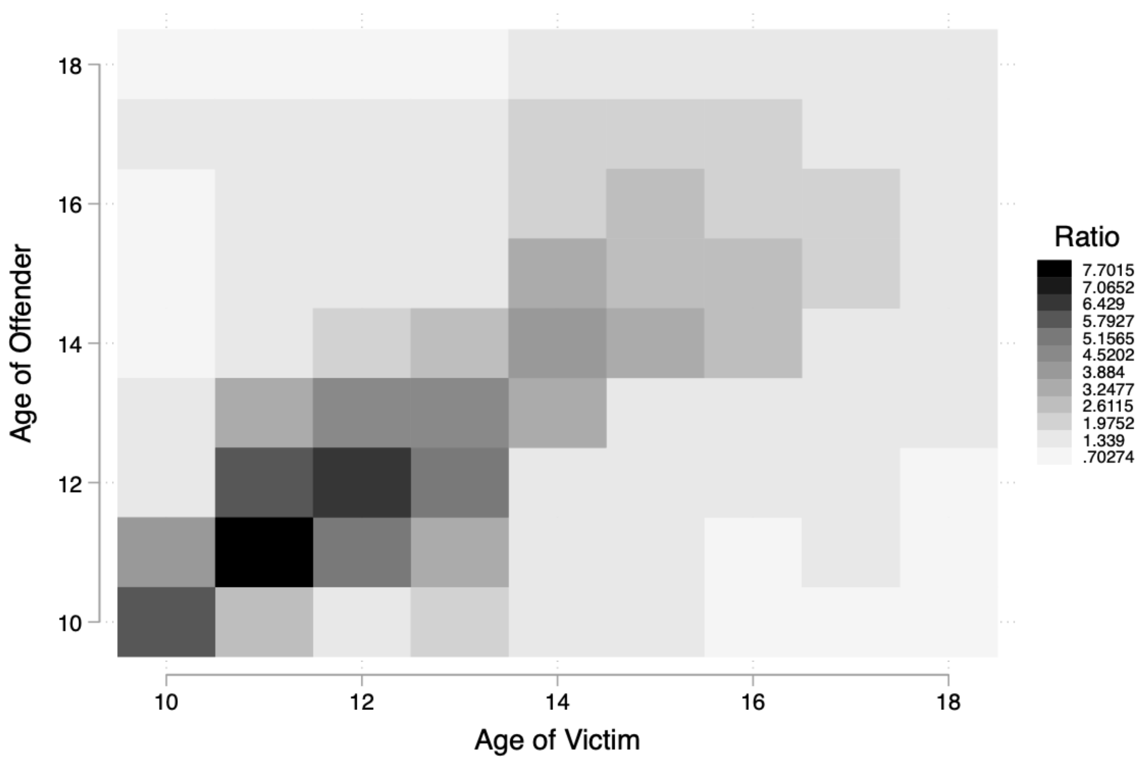
(e) Not in School; 25-30



(f) In School; 25-30

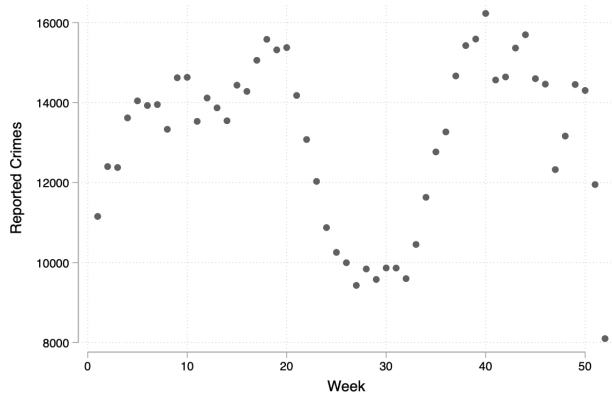
Notes: These figures show the total number of reported crimes by week for crimes *not* occurring in locations coded as "school - elementary/secondary" or "school/college" (Panels A, C, and E) and for crimes occurring in locations coded as "school - elementary/secondary" or "school/college" (Panels B, D, and E). Panels A and B are limited to the offender being 10-17 years old; Panels C and D (E and F) are limited to the offender being 19-24 (25-30) years old. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

Figure 7: Reported Crimes by Week, Fall vs. Summer, by Age of Offender and Age of Victim

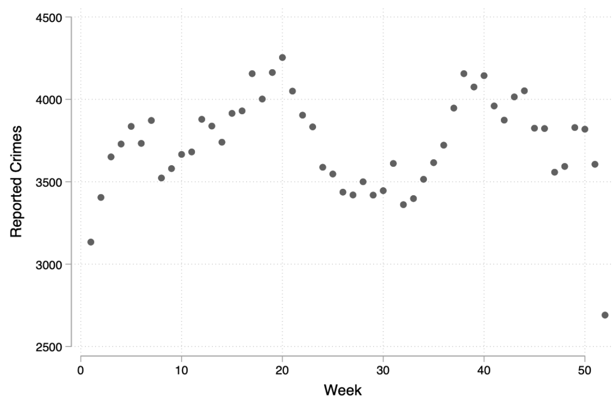


Notes: This figure shows, for each offender age (column)-victim age (row) pair, the number of crimes between weeks 37 and 40 (roughly the early fall peak) divided by the number of crimes between weeks between weeks 27 and 30 (roughly the summer trough). Darker shades represent larger relative changes in crime from this measure of trough vs. peak. The data source is 2017-2019 NIBRS.

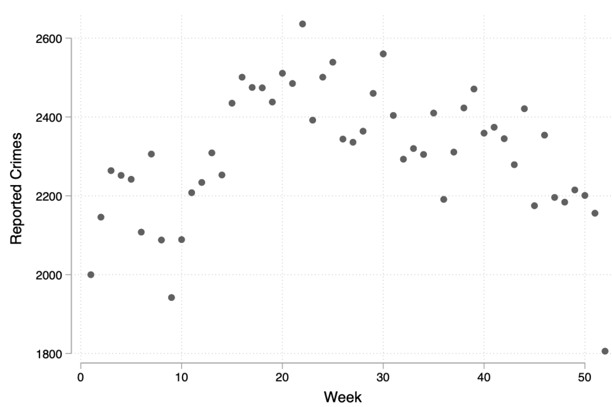
Figure 8: Reported Crimes by Week, Age 10–17, by Number of Offenders in Incident



(a) 1 Offender



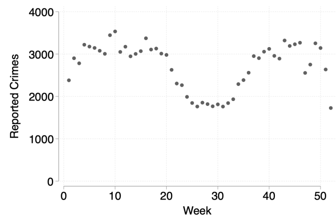
(b) 2 Offenders



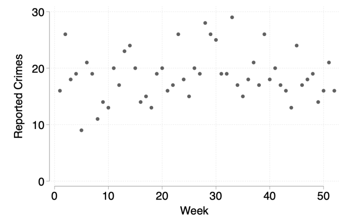
(c) 3 or More Offenders

Notes: These figures show the total number of reported crimes by week for incidents with 1 offender (Panel A), 2 offenders (Panel B), and 3 or more offenders (Panel C). The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

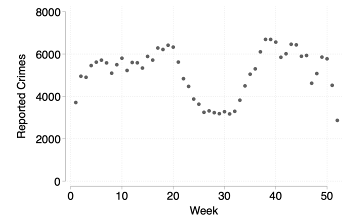
Figure 9: Reported Crimes by Week, Age 10-17, by Type of Offense



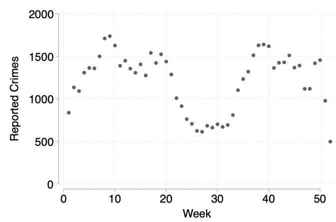
(a) Drug



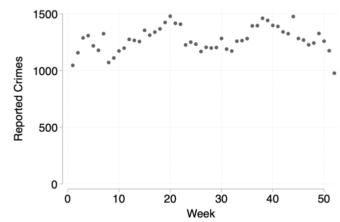
(b) Killing



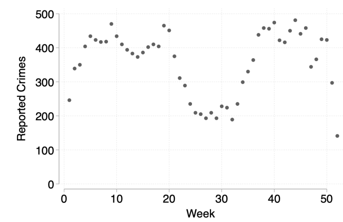
(c) Simple Assault



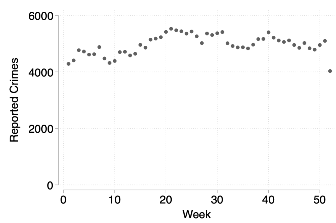
(d) Intimidation



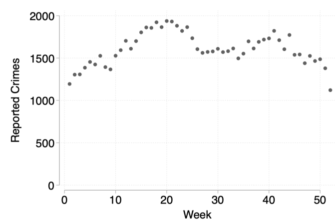
(e) Other Violent



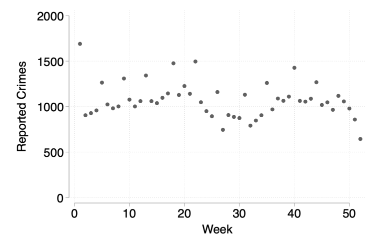
(f) Weapons



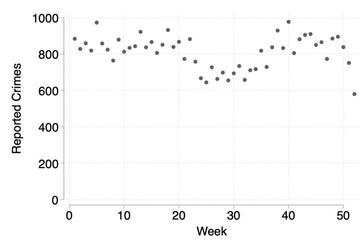
(g) Theft



(h) Property Damage



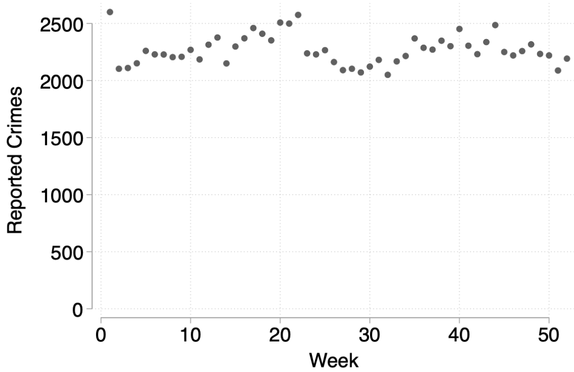
(i) Sexual Assault



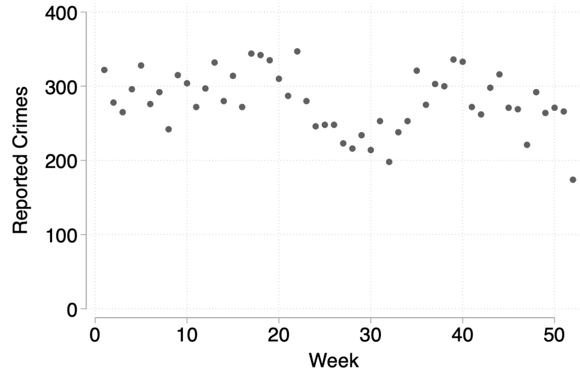
(j) Other Non-violent

Notes: These figures show the total number of reported crimes by week for offenders aged 10-17. Each panel corresponds to a different type of crime. Panel a) shows drug crimes; Panel b) shows killings; Panel c) shows simple assault; Panel d) shows intimidation; Panel e) shows other violent crime; Panel f) shows weapons; Panel g) shows theft; Panel h) shows property damage/vandalism; Panel i) shows sexual assault; and Panel j) shows other non-violent. Categories are as follows: Drug: drug equipment violations; drug/narcotic violations. Killing: justifiable homicide; murder/nonnegligent manslaughter; negligent manslaughter. Simple Assault: simple assault. Intimidation: intimidation. Other Violent: aggravated assault; kidnapping/abduction; robbery. Weapons: weapon law violations. Theft: all other larceny; burglary/breaking and entering; motor vehicle theft; pocket-picking; purse-snatching; shoplifting; theft from building; theft from coin-operated machine or device; theft from motor vehicle; theft of motor vehicle parts/accessories. Property Damage/Vandalism: destruction/damage/vandalism of property. Sexual Assault: fondling (incident liberties/child molest); incest; rape; sexual assault with an object; sodomy; statutory rape. Other Non-violent: animal cruelty; arson; assisting or promoting prostitution; betting/wagering; bribery; counterfeiting/forgery; credit card/ATM fraud; embezzlement; extortion/blackmail; false pretenses/swindle/confidence game; gambling equipment violations; hacking/computer invasion; human trafficking - commercial sex acts; human trafficking - involuntary servitude; identity theft; impersonation; operating/promoting/assisting gambling; pornography/obscene material; prostitution; purchasing prostitution; sports tampering; stolen property offenses (receiving, selling, etc.); welfare fraud; wire fraud. Unknown crime types are omitted. The data source is 2017-2019 NIBRS.

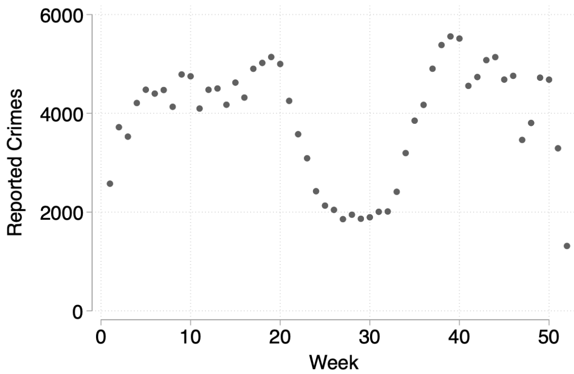
Figure 10: Reported Crimes by Week, Age 10-17, by Relationship of Offender to Victim



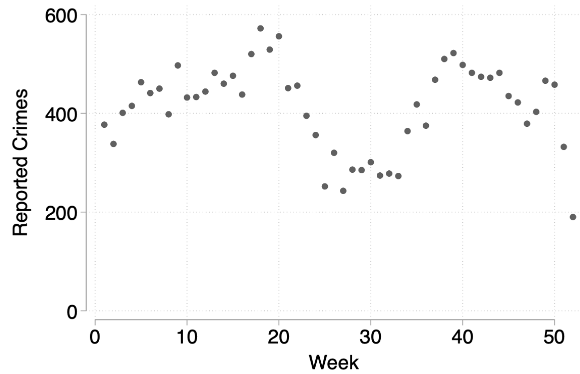
(a) Family



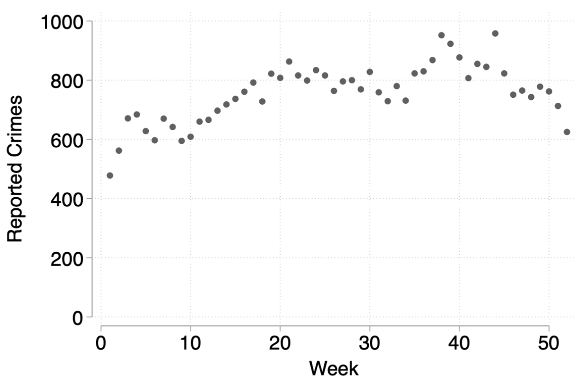
(b) Romantic



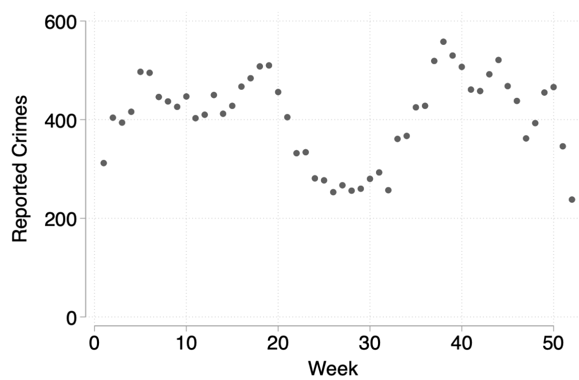
(c) Acquaintance



(d) Friend



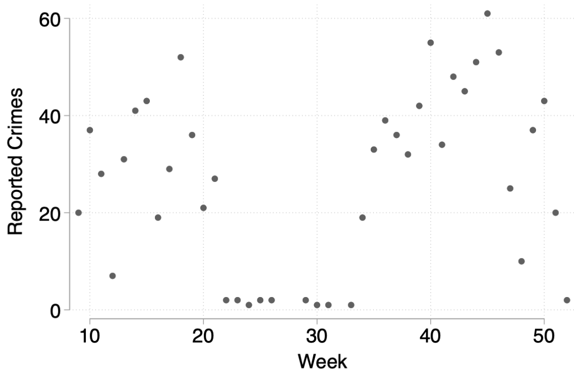
(e) Stranger



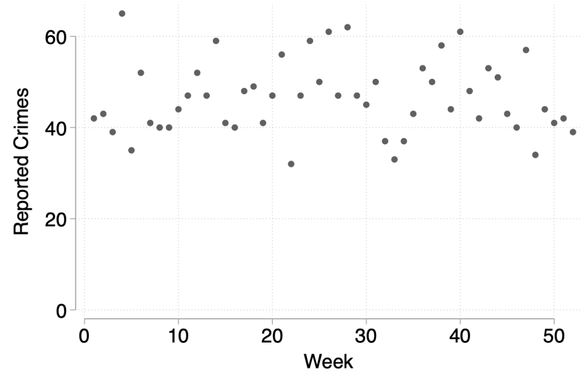
(f) Victim is Offender

Notes: These figures show the total number of reported crimes by week for ages 10-17. Each panel corresponds to the relationship of the first offender to the first victim. Family includes spouse, common-law spouse, parent, sibling, child, grandparent, grandchild, in-law, stepparent, stepchild, stepsibling, and other family member. Romantic includes boyfriend/girlfriend, and homosexual relationship. Acquaintance includes acquaintance, neighbor, babysittee (the baby), child of boyfriend/girlfriend, employee, employer, otherwise known, ex-spouse, and ex-relationship (boyfriend/girlfriend). Friend includes friend. Stranger includes stranger. Victim Is Offender includes victim was offender. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

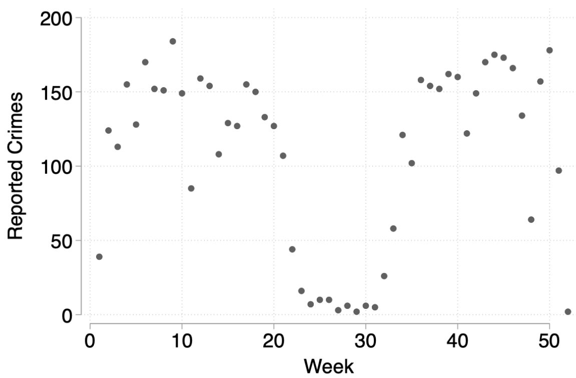
Figure 11: Reported Crimes by Week, School-based Police Agencies



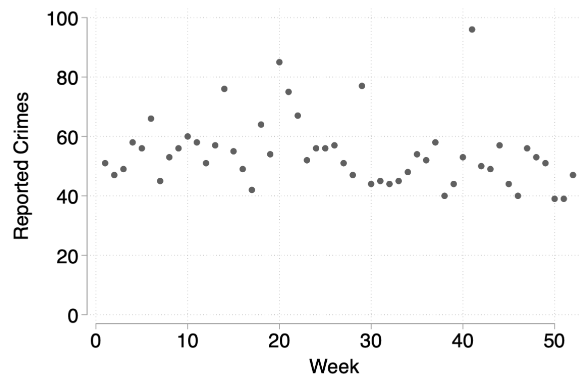
(a) Austin ISD Police Agency



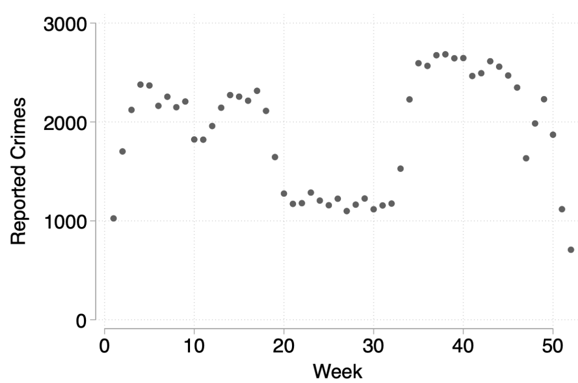
(b) Austin Police Agency



(c) School Police Agencies



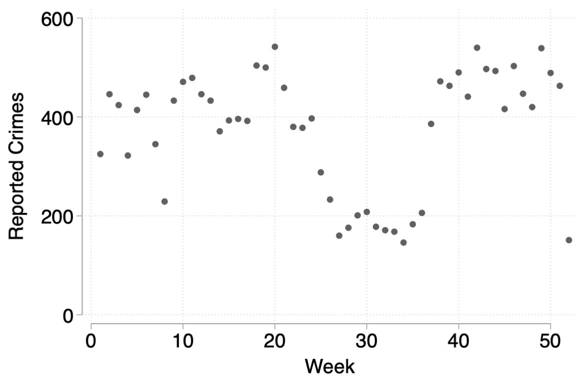
(d) Police Agencies in Same County as School Police Agencies



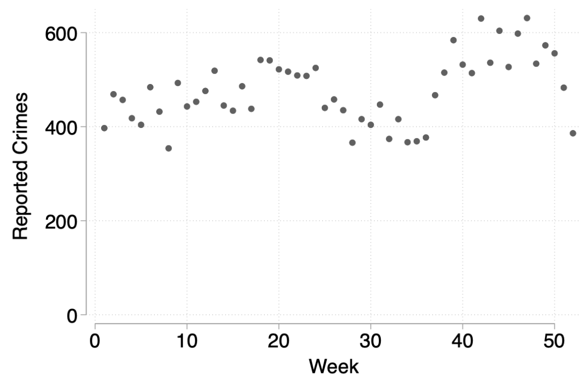
(e) College Police agencies

Notes: These figures show the total number of reported crimes by week for ages 10-17. Panel A is for to the Austin Independent School District police agency (which began reporting data partway through the year). Panel B is for the Austin police agency. Panel C is for school police agencies, identified by their name including “school(s)”, “ISD”, or “I.S.D.” Panel D is for police agencies in the same county as the school police agencies that begin with the same first word as the name of the school district(s) in the county, excluding police agencies with “college,” “university,” and “health science” (or one other university) in the title. Panel E is for college police agencies, identified by their name including “college” or “university.” Age is 10-17 for all panels except Panel E, which is unrestricted. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

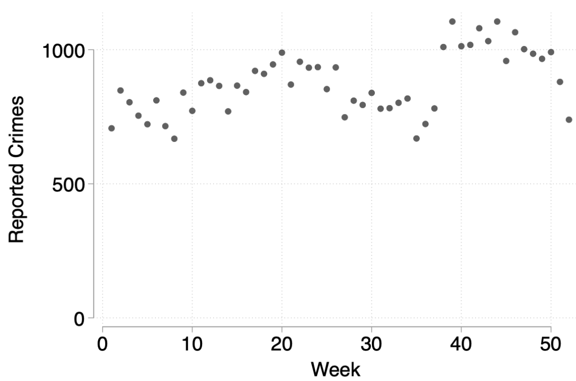
Figure 12: Reported Crimes by Week, Age 17 and Younger, New York City, Donut Hole by Distance to Nearest School



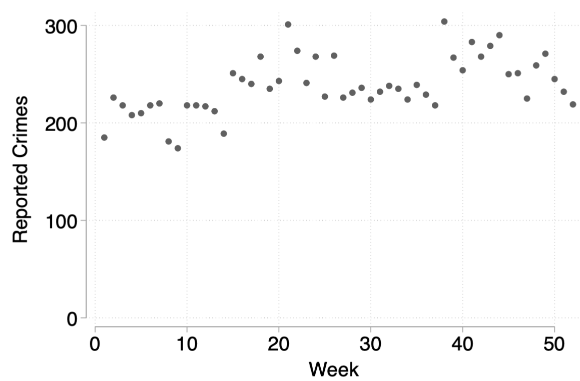
(a) <.05 Miles



(b) [.05, .1) Miles



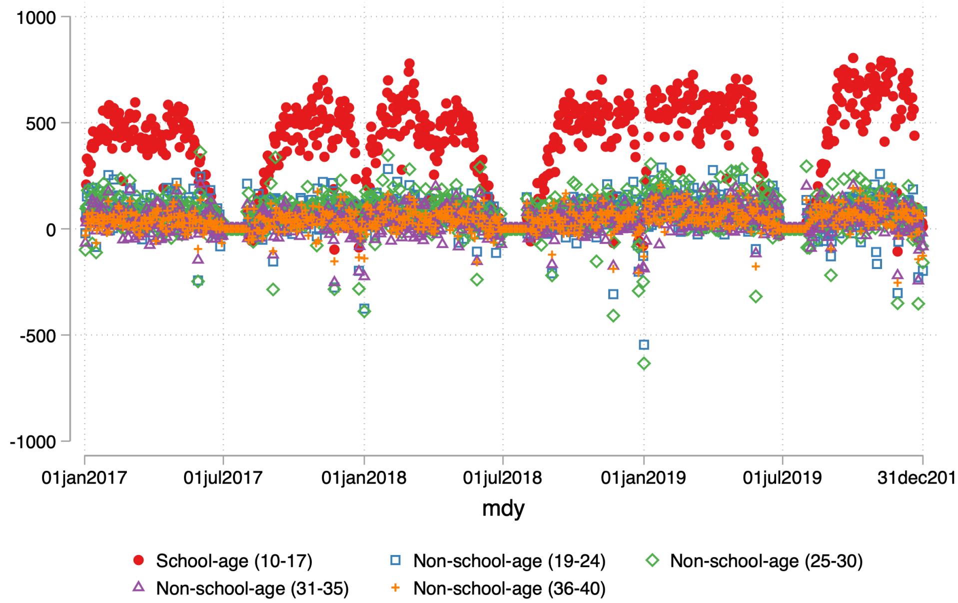
(c) [.1, .25) Miles



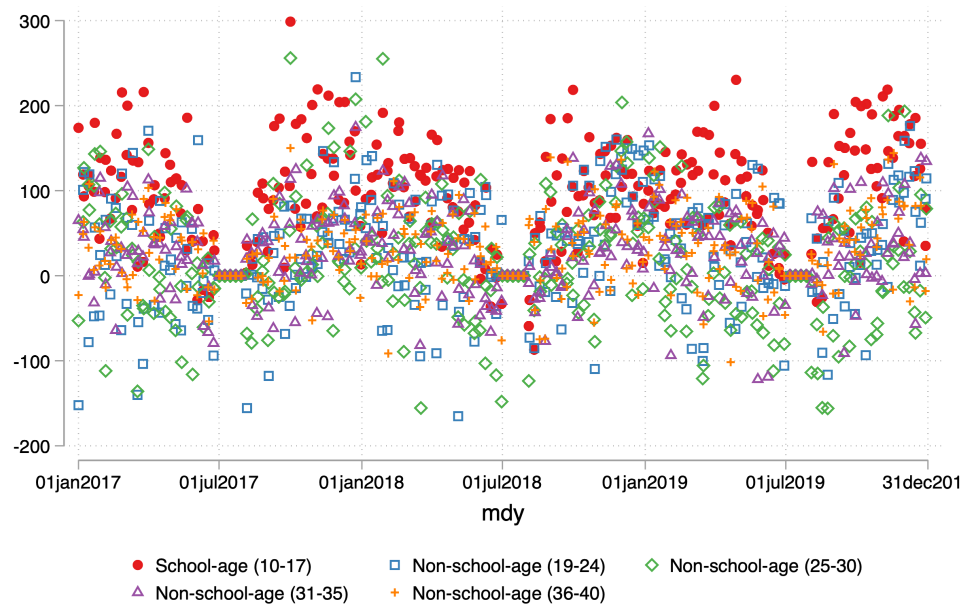
(d) .25+ Miles

Notes: These figures show the total number of reported crimes by week for ages 17 and younger using New York City data. Each observation is matched, using latitudes and longitudes, to the nearest school geocode using data from NCES. Each panel presents a graph limiting to reported crimes occurring in a certain distance range from the nearest school: less than 0.05 miles for Panel A; between .05 and .1 miles for Panel B; between 0.1 and 0.25 miles for Panel C; and 0.25 miles and greater for Panel D. Observations are limited to those occurring between 2006-2019. Week 1 starts on January 1.

Figure 13: Seasonally-adjusted Reported Crime



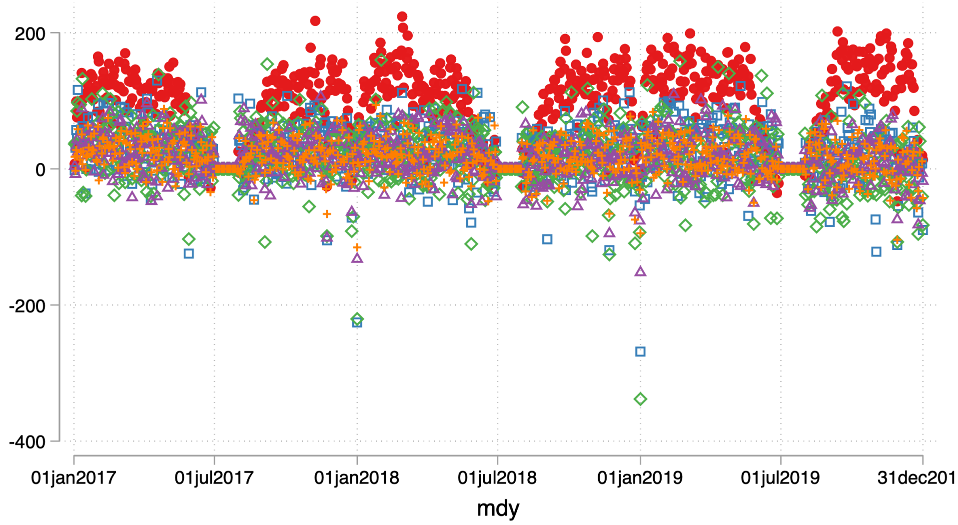
(a) Weekday Reported Crimes (residualized)



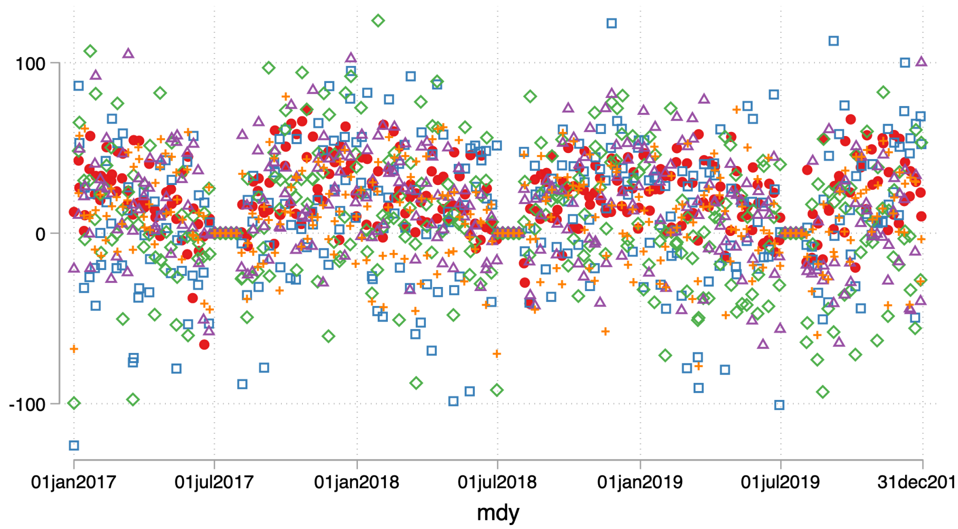
(b) Weekend Reported Crimes (residualized)

Notes: Seasonal adjustment calculated by regressing daily counts of reported crime in the 2017–2019 NIBRS panel (by calendar day by- hour of the day -by- age of offender) on calendar day -by- age fixed effects and day-of-the-week -by- hour-of-the-day -by- age fixed effects for crime reported outside of the school calendar and school hours (in July or between 8pm and 6am). Subtracting these fixed effects from reported cell-level crime and summing the resulting residual by calendar day produces the (daily) residuals presented above (pegged to July 1st). The top panel shows the residuals for weekdays, and the bottom panel shows residuals for weekends.

Figure 14: Seasonally-adjusted Reported Arrests



(a) Weekday Reported Arrests (residualized)



(b) Weekend Reported Arrests (residualized)

Notes: Seasonal adjustment calculated by regressing daily counts of arrests in the 2017–2019 NIBRS panel (by calendar day by- hour of the day -by- age of offender) on calendar day -by- age fixed effects and day-of-the-week -by- hour-of-the-day -by- age fixed effects for arrests reported outside of the school calendar and school hours (in July or between 8pm and 6am). Subtracting these fixed effects from reported cell-level arrests and summing the resulting residual by calendar day produces the (daily) residuals presented above (pegged to July 1st). The top panel shows the residuals for weekdays, and the bottom panel shows residuals for weekends.

Table 1: Excess Crime

Panel A: All Incidents	Excess Crime	% Increase
All	0.32	0.47
White	0.32	0.46
Black	0.32	0.47
Male	0.32	0.47
Female	0.32	0.46
Panel B: Arrests	Excess Crime	% Increase
All	0.29	0.41
White	0.29	0.41
Black	0.29	0.41
Male	0.30	0.42
Female	0.28	0.39

This table reports how much 'excess crime' there is for 10–17 year olds, for all incidents in Panel A, and for arrests in Panel B. Each line corresponds to a different sample. To obtain the estimates, we perform a seasonal adjustment calculated by regressing daily counts of reported crime in the 2017–2019 NIBRS panel (by calendar day by- hour of the day -by- age of offender) on calendar day -by- age fixed effects and day-of-the-week -by- hour-of-the-day -by- age fixed effects for crime reported outside of the school calendar and school hours (in July or between 8pm and 6am). Subtracting these fixed effects from reported cell-level crime and summing the resulting residual by calendar day for produces the (daily) residuals for each age group. We then compute the average residual for each age group. We subtract the mean from the control group, 19–24 year olds, after scaling it to account for differences in sample size between it and the 10–17 year old group. We multiply this object by 3*365 and to account for the number of days in the three years (call this Z), which we divide by the total number of crimes for the 10–17 year old group (call this N) to obtain the number in the 'Excess Crime' column. The '% Increase' column is obtained by dividing Z by (N minus Z).

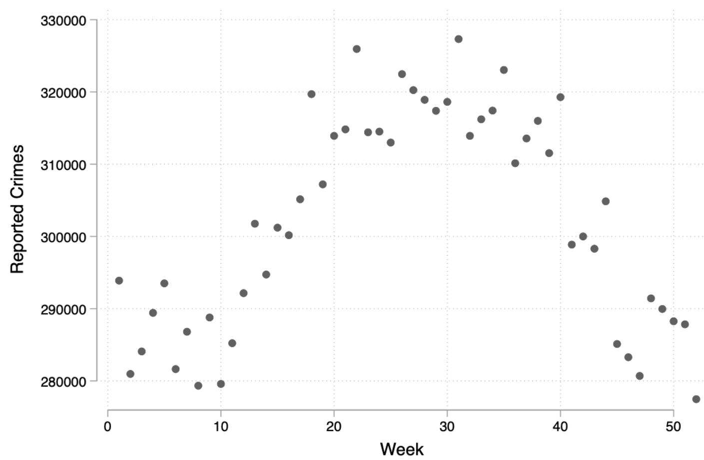
Table 2: County Excess Crime Predictors, Weight by County Population

Poverty %	0.00	0.05	0.05	0.05
	(0.04)	(0.05)	(0.05)	(0.04)
Unemployment Rate	-0.01	-0.03	-0.03	-0.03
	(0.03)	(0.04)	(0.04)	(0.05)
Median HH Income	-0.04	-0.04	-0.04	-0.04
	(0.03)	(0.03)	(0.03)	(0.03)
Population	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.00)
% Less than High School	-0.05	-0.02	-0.02	-0.02
	(0.04)	(0.04)	(0.04)	(0.05)
Rural [0,1]	0.05	0.05	0.05	0.05
	(0.07)	(0.07)	(0.07)	(0.06)
Expenditures per Student	-0.01	-0.01	-0.01	-0.01
	(0.02)	(0.04)	(0.04)	(0.04)
Students per Grade	0.08	0.10*	0.10*	0.10
	(0.05)	(0.05)	(0.05)	(0.06)
% White and Asian	-0.04	-0.00	-0.00	-0.00
	(0.03)	(0.04)	(0.04)	(0.04)
Mixed HS Jr High	-0.10*	-0.07	-0.07	-0.07
	(0.06)	(0.07)	(0.07)	(0.07)
Any SRO or Security Guard	-0.11**	-0.04	-0.04	-0.04
	(0.05)	(0.04)	(0.04)	(0.07)
Y Mean	0.33	0.33	0.33	0.33
N	1,824	1,820	1,820	1,820
R-squared	0.02	0.08	0.08	0.08
State FEs	No	Yes	Yes	Yes
SEs			Robust	State Cluster

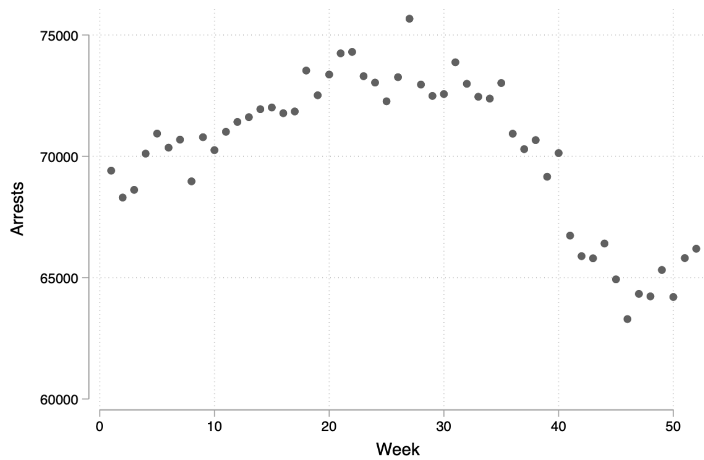
This table reports a regression of county-level excess crime on a number of variables. It weights counties by population. Unless indicated by '[0,1]', each variable is standardized (among the 1,824 counties) to have mean 0 and standard deviation 1. The first column does not include fixed effects or adjustments for standard errors. Columns 2–4 include state fixed effects. Column 3 uses robust standard errors. Column 4 clusters the standard errors at the state level. * <.1; ** <.05; *** <.01.

A Appendix Figures

Figure A.1: Reported Crimes and Arrests by Week, Full Sample



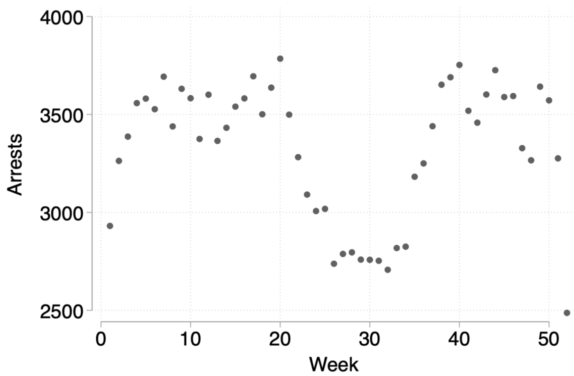
(a) All Incidents



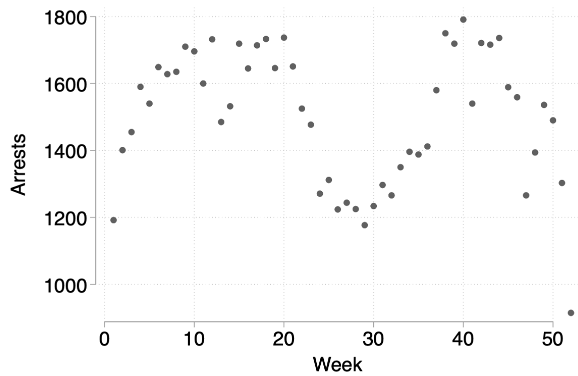
(b) Arrests

Notes: These figures show the total number of reported crimes and arrests by week. Panel A shows the total number of reported crimes, and Panels B shows the total number of arrests and times taken into custody. The data source is 2017–2019 NIBRS. Week 1 starts on January 1.

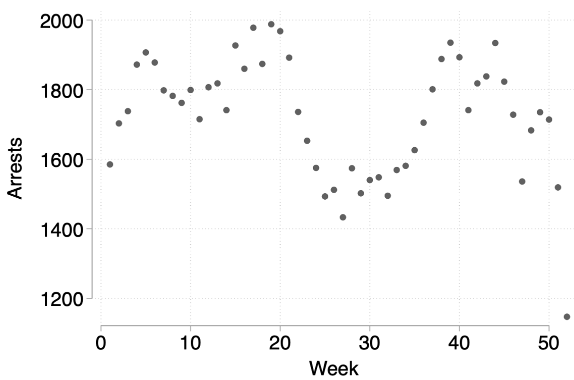
Figure A.2: Arrests by Week, by Type of Arrest, Group A and Group B Arrests



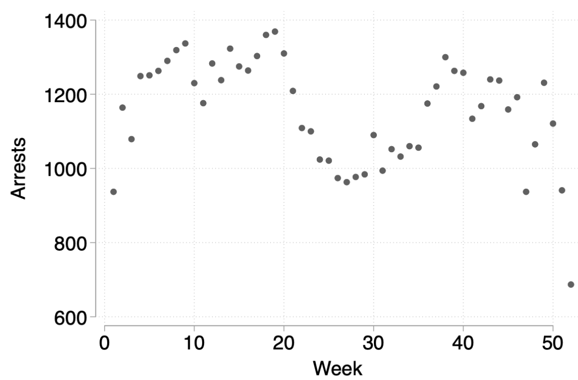
(a) Group A, On-view Arrest



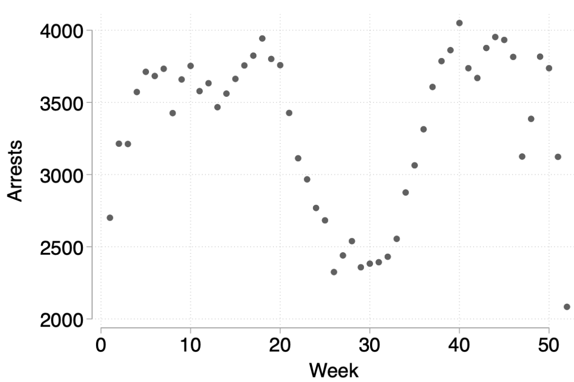
(b) Group B, On-view Arrest



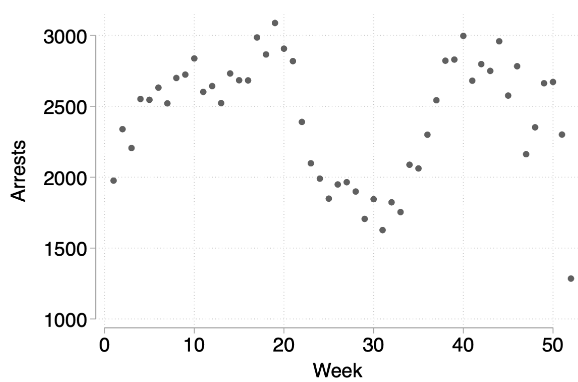
(c) Group A, Warrant/Prior Incident



(d) Group B, Warrant/Prior Incident



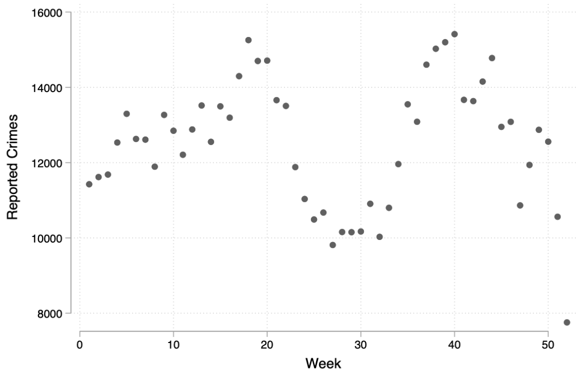
(e) Group A, Summoned/Cited



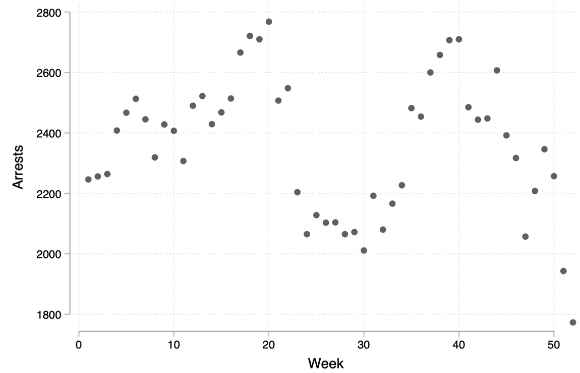
(f) Group B, Summoned/Cited

Notes: These figures show the total number of arrests by week for 10-17 year olds. Panels A, C, and E are for Group A arrests, while Panels B, D, and F are for Group B arrests. Panels A and B show arrests in the "on-view arrest (taken into custody without a warrant or previous incident report)" category; Panels C and D show arrests in the "taken into custody (based on warrant and/or previous incident report)" category; and Panels E and F shows arrests in the "summoned/cited (not taken into custody)" category. The data source is 2017–2019 NIBRS. Week 1 starts on January 1.

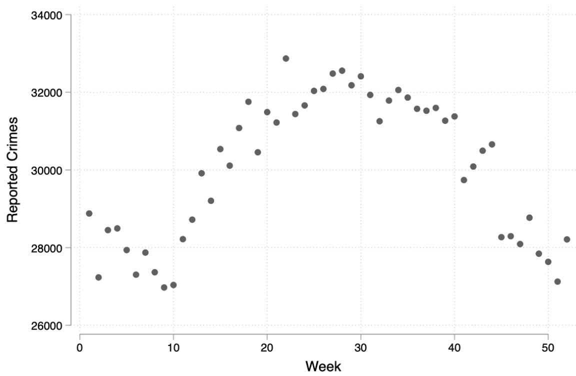
Figure A.3: Reported Crimes and Arrests by Week, By Age of Victim



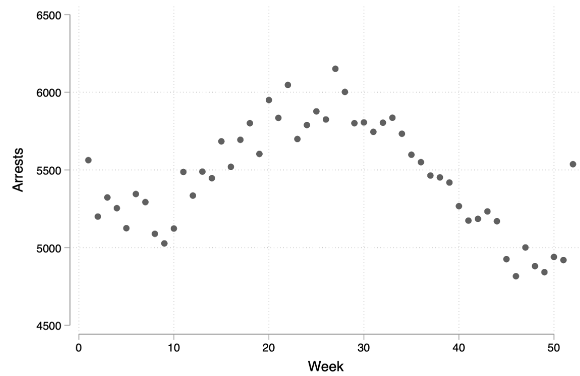
(a) Victim Age 10-17, All Incidents



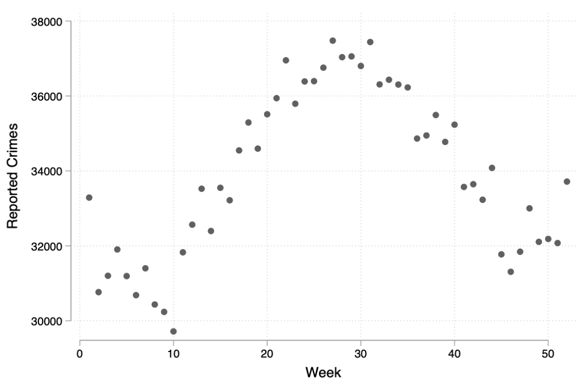
(b) Victim Age 10-17, Arrests



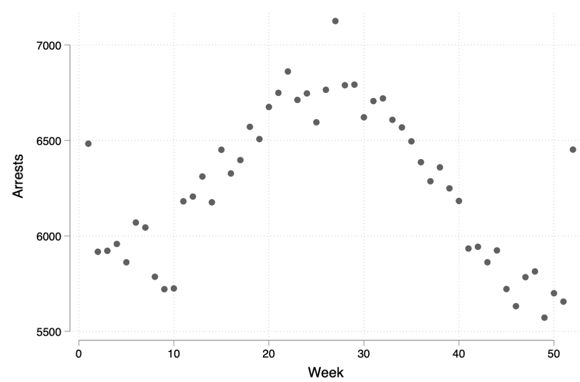
(c) Victim Age 19-24, All Incidents



(d) Victim Age 19-24, Arrests



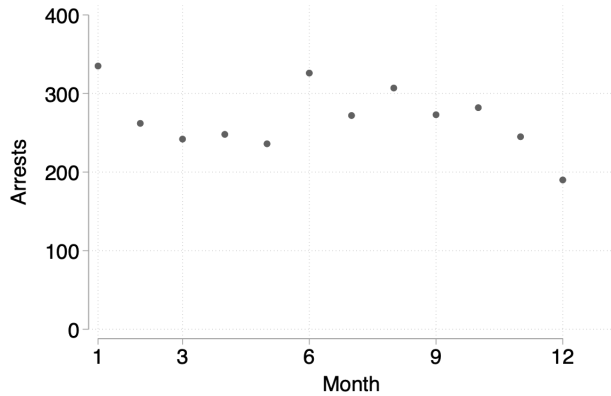
(e) Victim Age 25-30, All Incidents



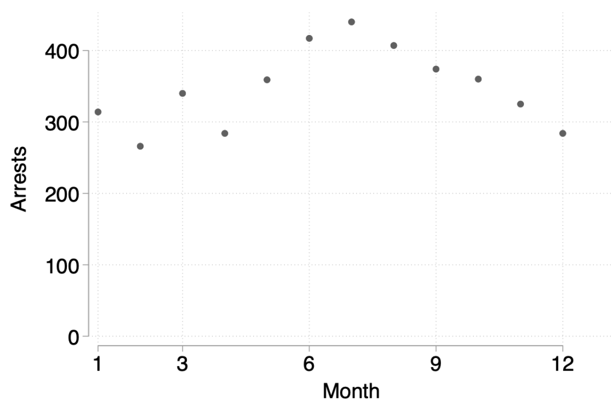
(f) Victim Age 25-30, Arrests

Notes: These figures show the total number of reported crimes and arrests by week by the age of the victim. Panels A, C, and E show the total number of reported crimes, and Panels B, D, and F show the total number of arrests. Panels A and B are limited to the victim being 10-17 years old; Panels C and D (E and F) are limited to the victim being 19-24 (25-30) years old. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

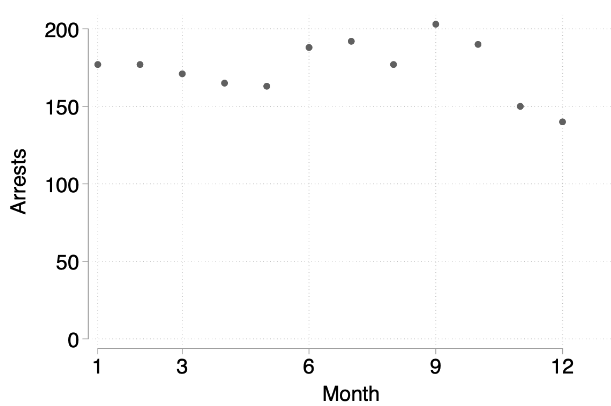
Figure A.4: Arrests by Month, NLSY97



(a) 10-17, Arrests



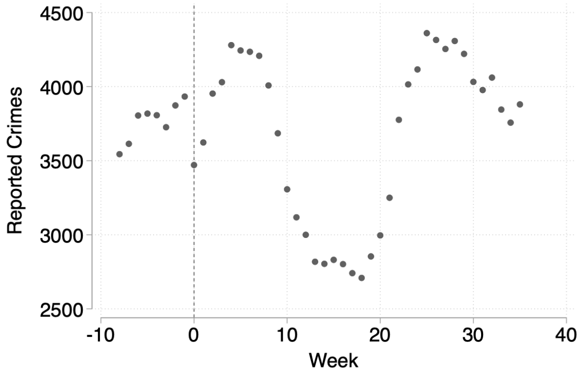
(b) 19-24, Arrests



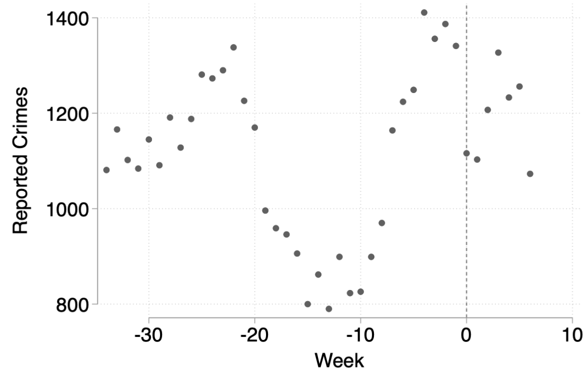
(c) 25-30, Arrests

Notes: These figures show the total number of arrests by week from the NSLY97 between 1992 and 2019. Panel A is limited to the arrestee being 10-17 years old; Panel B is limited to the arrestee being 19-24 years old; and Panel C is limited to the arrestee being 25-30 years old. Ages are calculated by subtracting the year and month of arrest by the year and month of birth.

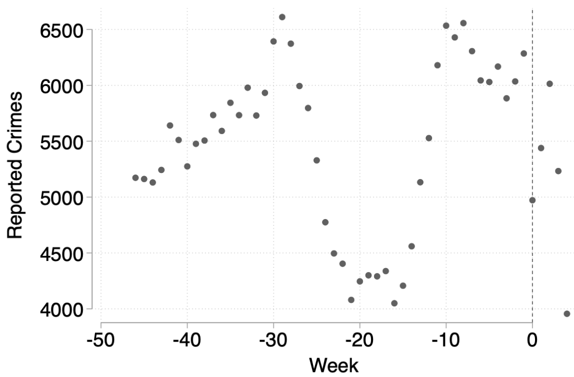
Figure A.5: Crimes by Week Relative to School District Start Date, by Type of Break, 10-17 Years



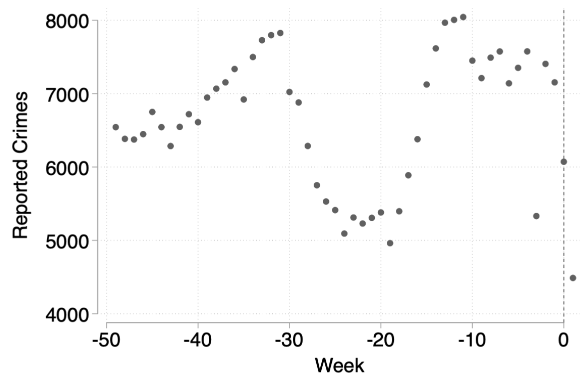
(a) Spring Break



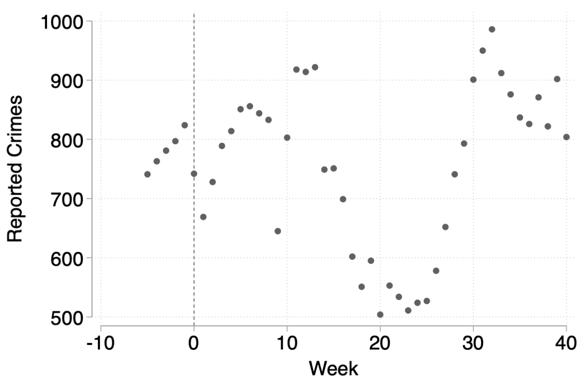
(b) Fall Break



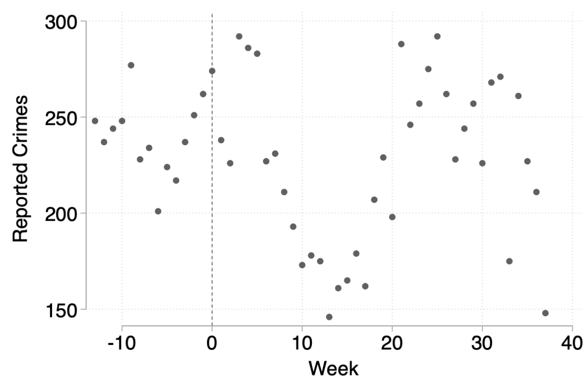
(c) Thanksgiving



(d) Christmas



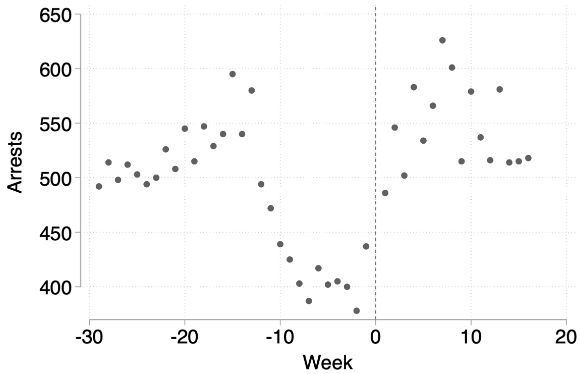
(e) Midwinter



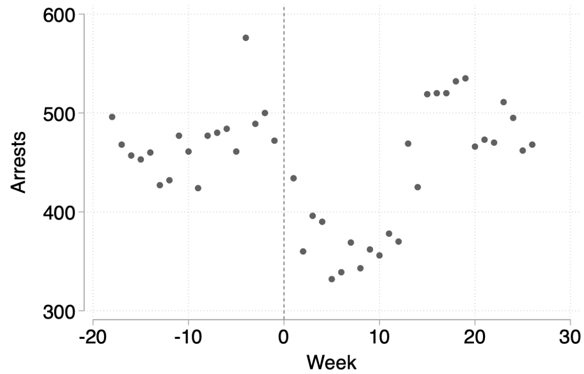
(f) Easter

Notes: These figures show a regression discontinuity for the number of crimes by week relative to school district start date for offenders 10-17 years old. Panel A is for spring break; Panel B is for fall break; Panel C is for Thanksgiving; Panel D is for Christmas; Panel E is for midwinter break; and Panel F is for Easter. The data source is 2019 NIBRS. Week 0 is the week in which the holiday break begins; this week is not excluded. A line is drawn at week 0.

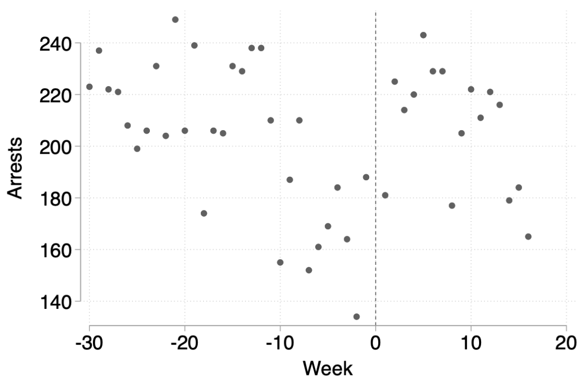
Figure A.6: Arrests by Week Relative to School Start and End Date, 10-17 Year Old, by Type of Arrest



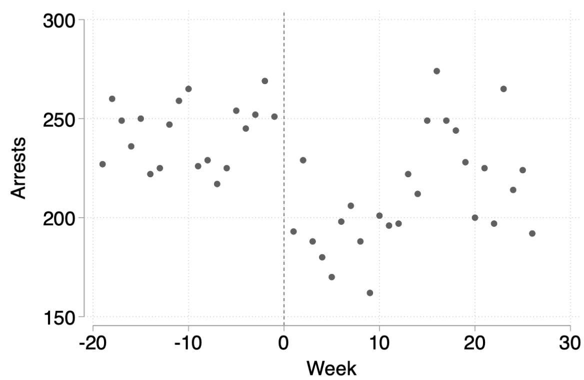
(a) School Start, On-view Arrest



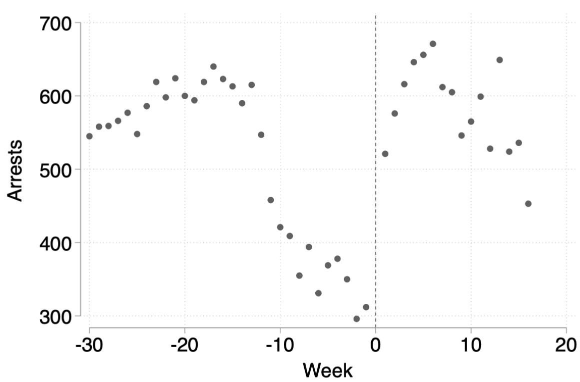
(b) School End, On-view Arrest



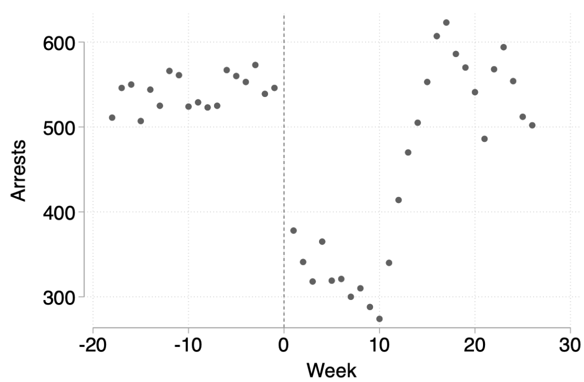
(c) School Start, Warrant/Prior Incident



(d) School End, Warrant/Prior Incident



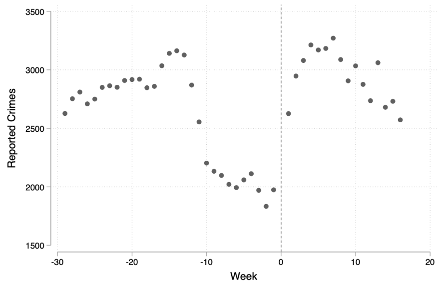
(e) School Start, Summoned/Cited



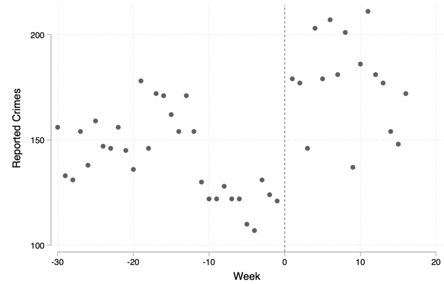
(f) School End, Summoned/Cited

Notes: These figures show a regression discontinuity for the number of arrests for offenders 10-17 years old by week relative to school district start date (Panels A, C, and E) and school end date (Panels B, D, and F). Panels A and B show arrests in the "on-view arrest (taken into custody without a warrant or previous incident report)" category; Panels C and D show arrests in the "taken into custody (based on warrant and/or previous incident report)" category; and Panels E and F shows arrests in the "summoned/cited (not taken into custody)" category. The data sources is 2019 NIBRS data. Week 0 is the school district starting or ending date. Week 0 is excluded as this is sometimes a partially treated week. A line is drawn at week 0.

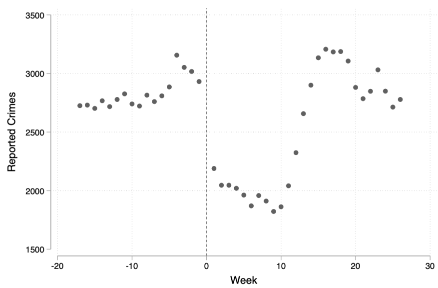
Figure A.7: All Incidents and Arrests by Week Relative to School Start and End Date, 10-17 Year Old, by Type of Date



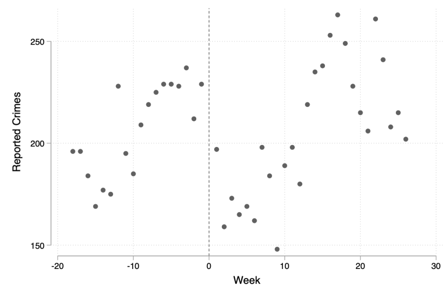
(a) All, School Start, Incident Date



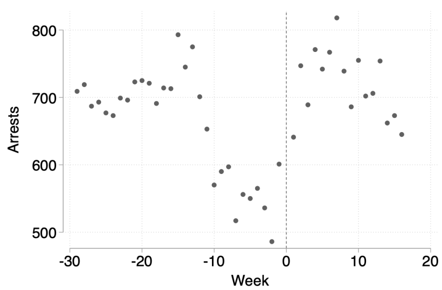
(b) All, School Start, Reported Date



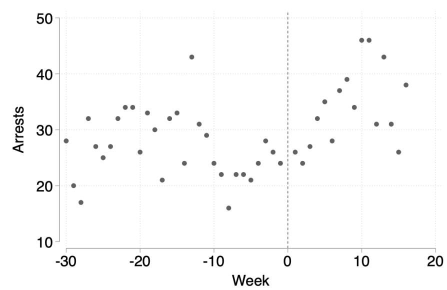
(c) All, School End, Incident Date



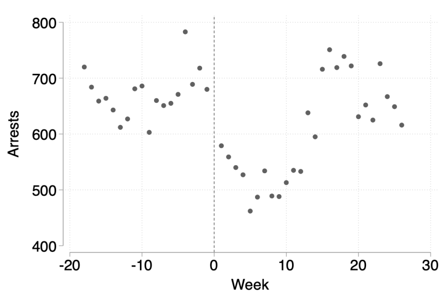
(d) All, School End, Reported Date



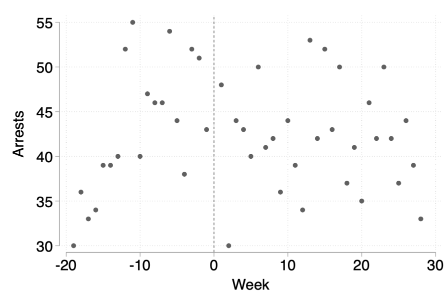
(e) Arrests, School Start, Incident Date



(f) Arrests, School Start, Reported Date



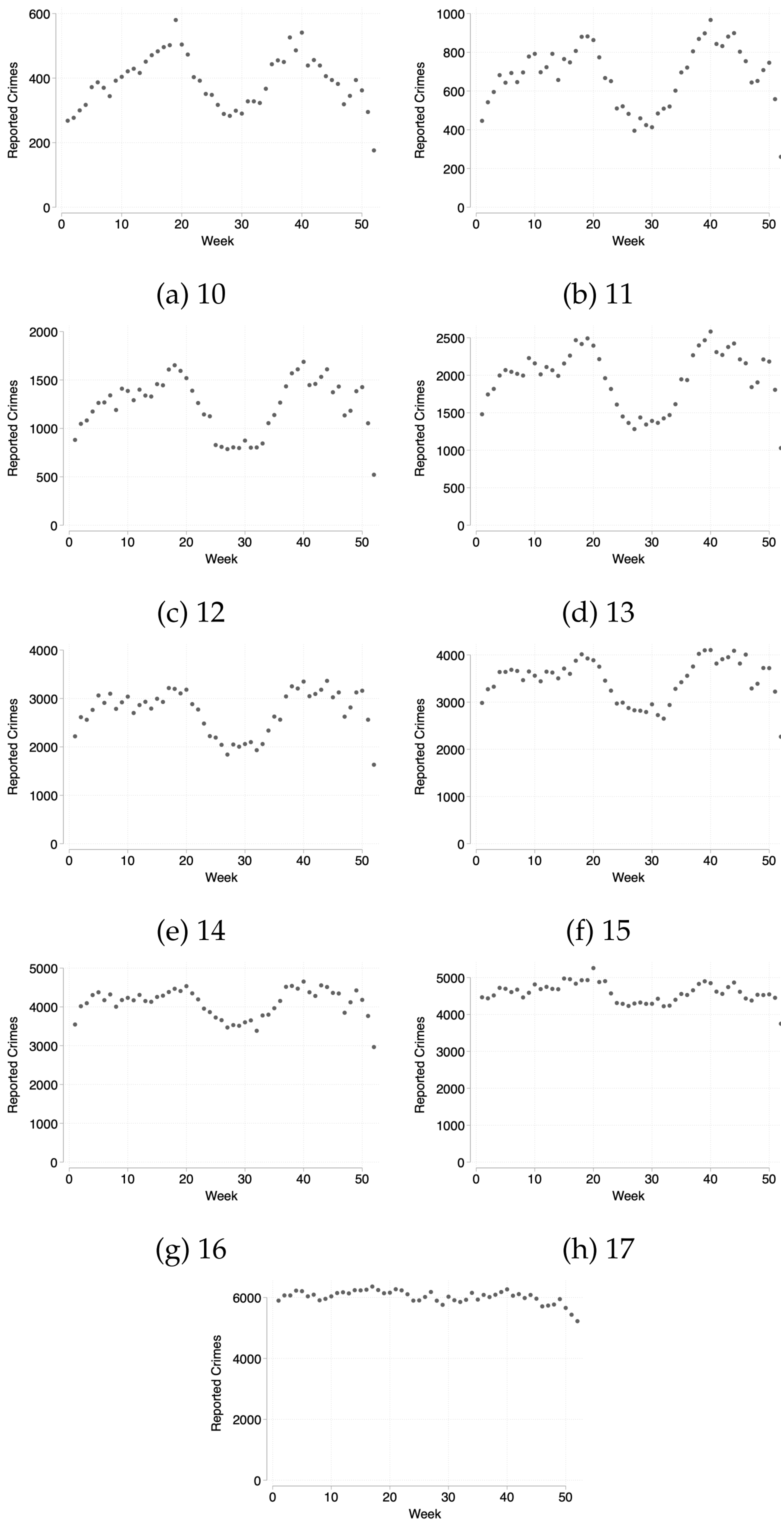
(g) Arrests, School End, Incident Date



(h) Arrests, School End, Reported Date

Notes: These figures show a regression discontinuity for the number of all incidents and the number of arrests for offenders 10-17 years old by week relative to school district start date (Panels A, B, E, and F) and school end date (Panels C, D, G, and H). Panels A, C, E, and G are not based on the date the incident was reported, while Panels B, D, F, and H are based on the date the incident was reported. The data source is 2019 NIBRS. Week 0 is the school district starting or ending date. Week 0 is excluded as this is sometimes a partially treated week. A line is drawn at week 0.

Figure A.8: Reported Crimes by Week, by Age of Offender



(a) 10

(b) 11

(c) 12

(d) 13

(e) 14

(f) 15

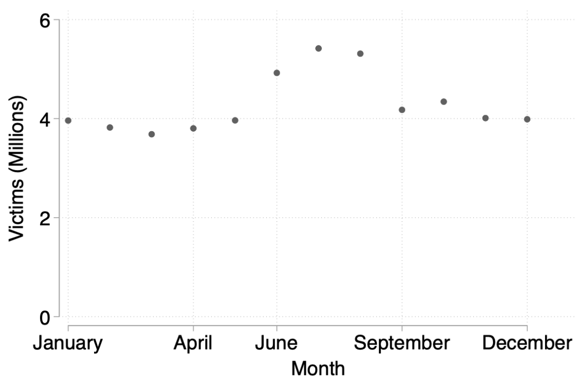
(g) 16

(h) 17

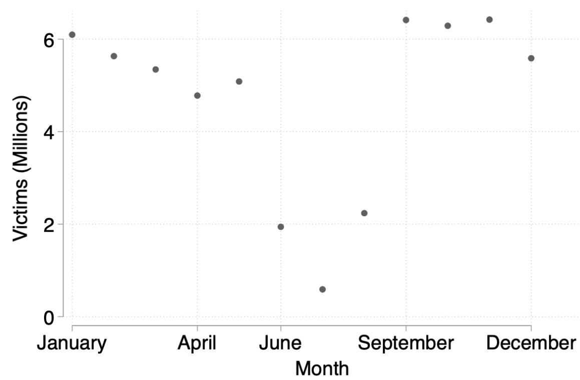
(i) 18

Notes: These figures show the total number of reported crimes by week. Each panel corresponds to a different age of offender (from age 10 to age 18). The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

Figure A.9: Reported Crimes by Month, NCVS, Age 10-17, by Location of Incident



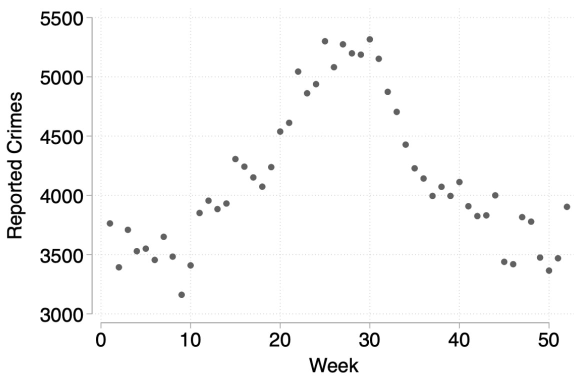
(a) Not in School



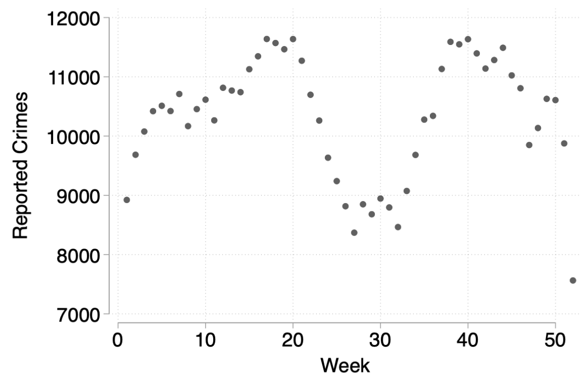
(b) In School

Notes: These figures show the total number (after weighting using person weights) of reported victimizations by week for victims aged 10-17. Panel A corresponds to incidents not in school, while Panel B corresponds to incidents in school. The data source is 1992-2019 National Crime Victimization Survey.

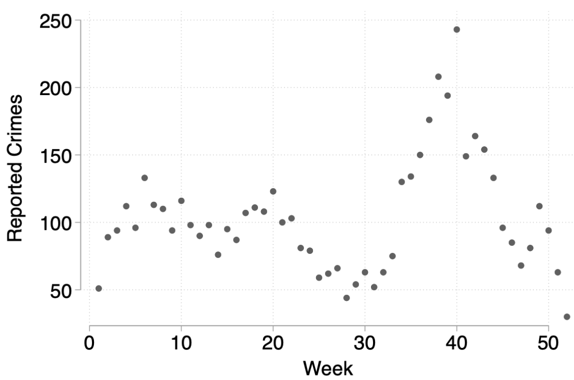
Figure A.10: Reported Crimes by Week, Age 10-17, by Location and Time



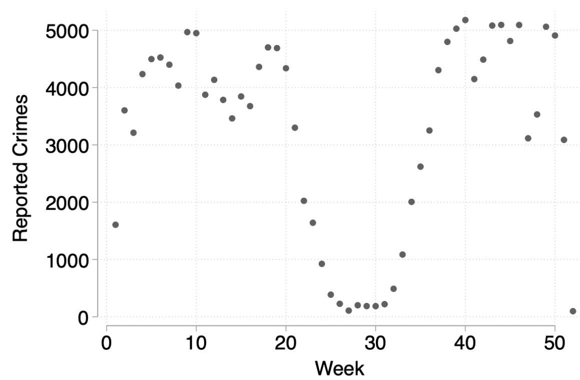
(a) Not in School; Not School Hours



(b) Not in School; School Hours



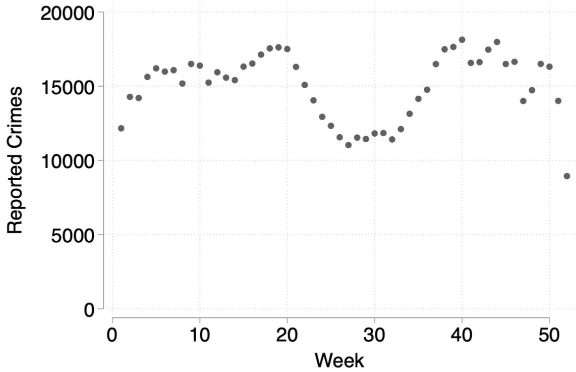
(c) In School; Not School Hours



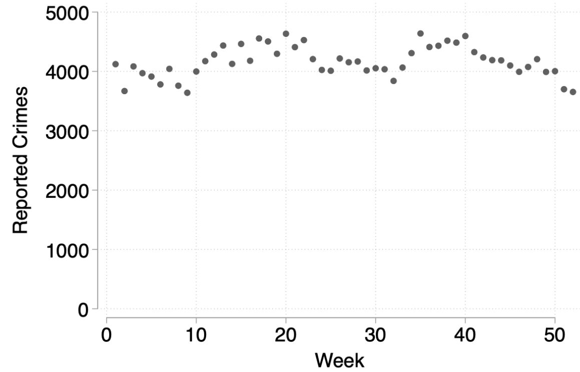
(d) In School; School Hours

Notes: These figures show, for offenders age 10–17, the total number of reported crimes by week for crimes *not* occurring in locations coded as "school - elementary/secondary" or "school/college" (Panels A and B) and for crimes occurring in locations coded as "school - elementary/secondary" or "school/college" (Panels C and D). Panels A and C restrict to non-school hours, defined as between 1am and 5:59am and between 8pm and 11:59pm. Panels B and D restrict to school hours, defined as between 7am and 7:59pm. Midnight to 12:59am and 6am to 6:59am are excluded. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

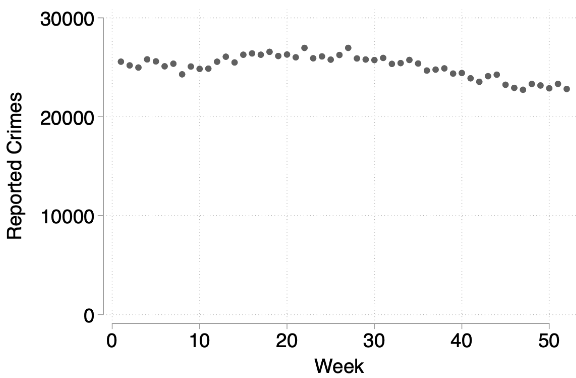
Figure A.11: Reported Crimes by Week, by Weekdays vs. Weekend



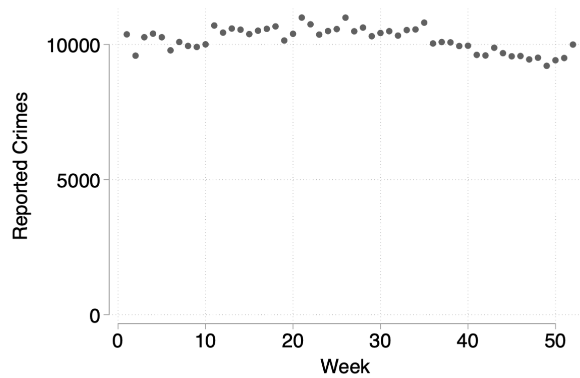
(a) Weekdays; 10-17



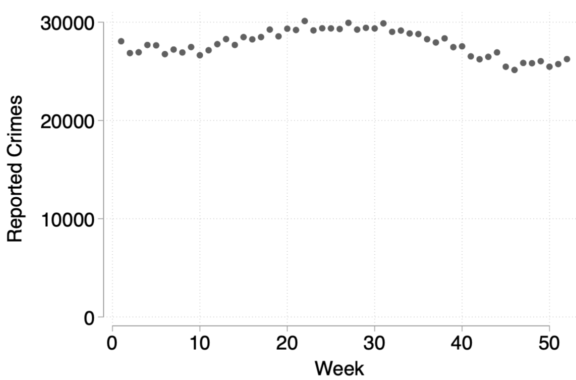
(b) Weekend; 10-17



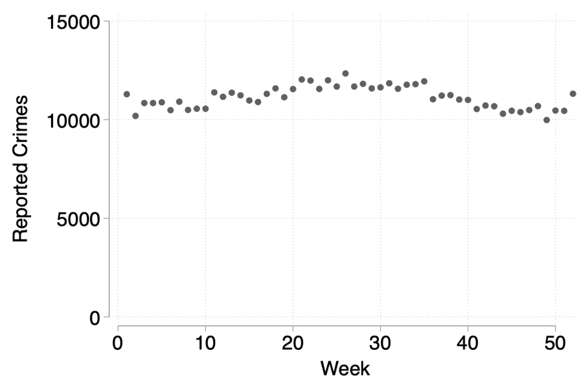
(c) Weekdays; 19-24



(d) Weekend; 19-24



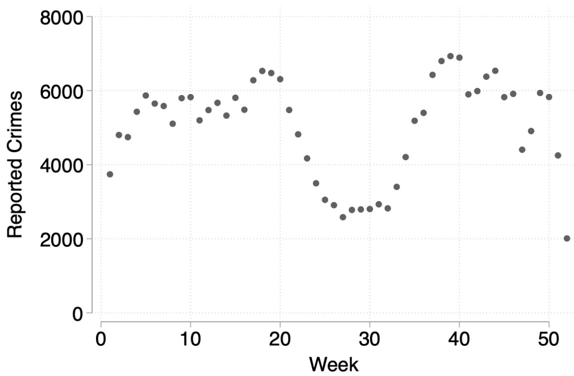
(e) Weekdays; 25-30



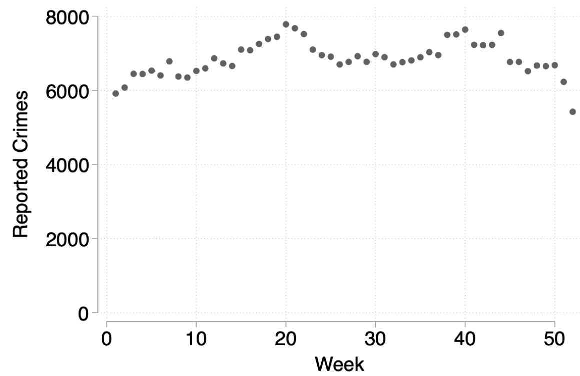
(f) Weekend; 25-30

Notes: These figures show the total number of reported crimes by week for crimes not occurring on weekdays (Monday through Friday; Panels A, C, and E) and for crimes occurring on weekends (Saturday and Sunday; Panels B, D, and E). Panels A and B are limited to the offender being 10-17 years old; Panels C and D (E and F) are limited to the offender being 19-24 (25-30) years old. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

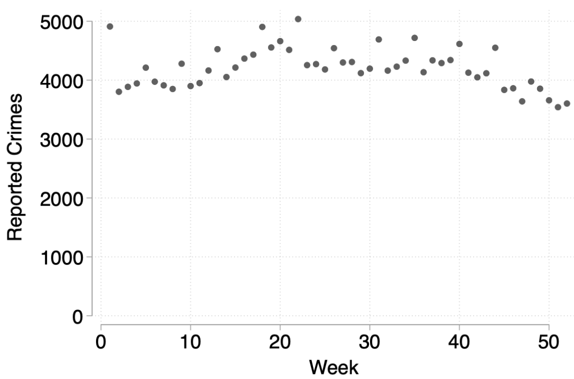
Figure A.12: Reported Crimes by Week, by Offender and Victim Age Categories



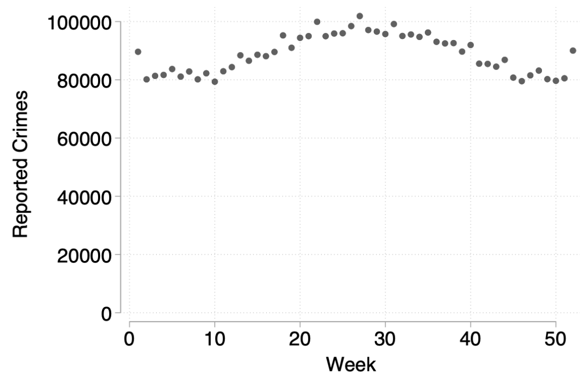
(a) 10-17 Offender; 10-17 Victim



(b) 10-17 Offender; 19+ Victim



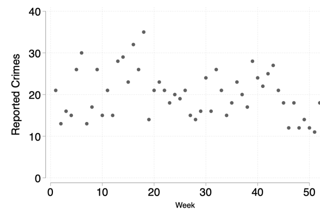
(c) 19+ Offender; 10-17 Victim



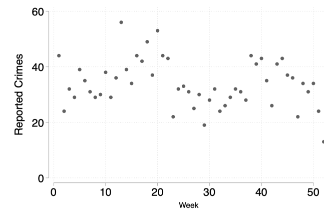
(d) 19+ Offender; 19+ Victim

Notes: These figures show the total number of reported crimes by week for crimes when the sample limited to 10-17 year old offender and 10-17 year old victim (Panel A); 10-17 year old offender and 19+ year old victim (Panel B); 19+ year old offender and 10-17 year old victim (Panel C); and 19+ year old offender and 19+ year old victim (Panel D). The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

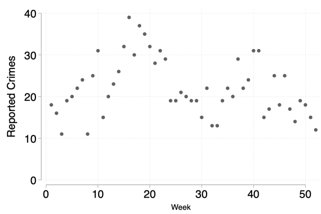
Figure A.13: Reported Crimes by Week, by Age of Offender and Age of Victim, Offenders Aged 12-14



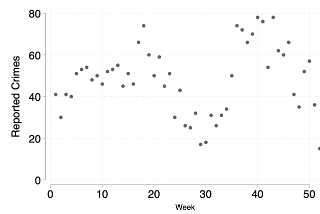
(a) O:13, V:10



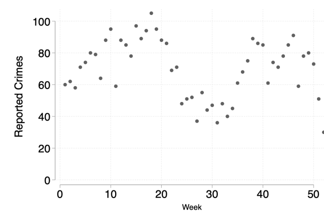
(b) O:14, V:11



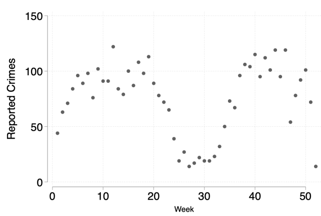
(c) O:12, V:10



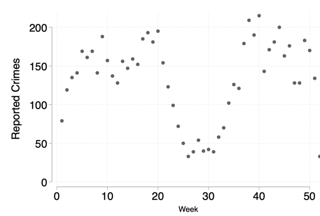
(d) O:13, V:11



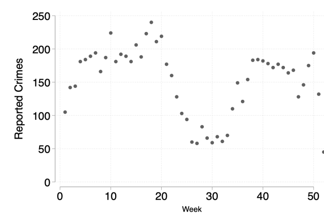
(e) O:14, V:12



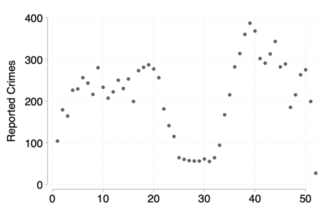
(f) O:12, V:11



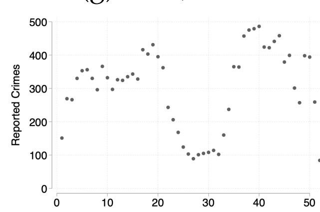
(g) O:13, V:12



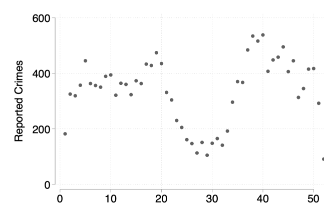
(h) O:14, V:13



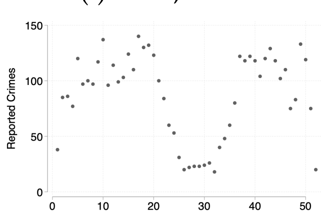
(i) O:12, V:12



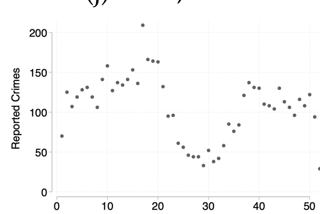
(j) O:13, V:13



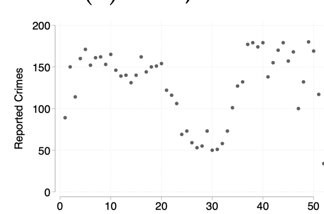
(k) O:14, V:14



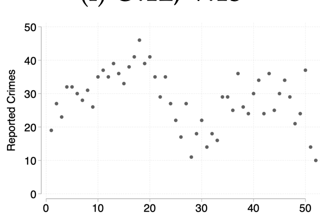
(l) O:12, V:13



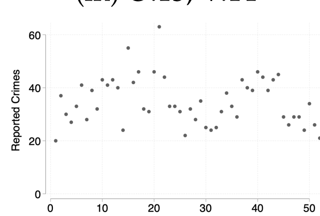
(m) O:13, V:14



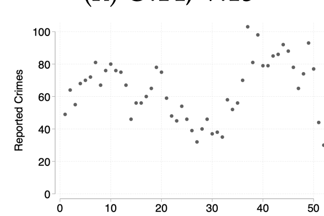
(n) O:14, V:15



(o) O:12, V:14



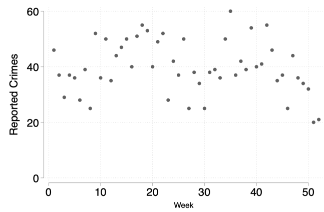
(p) O:13, V:15



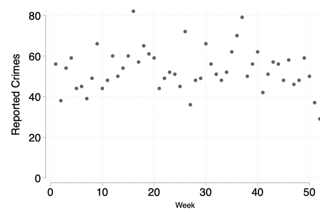
(q) O:14, V:16

Notes: These figures show the total number of reported crimes by week for age of first offender (O) and age of first victim (V) pairs. For example, "O: 13, V: 12" means that the offender was 13 and the victim was 12. Offender ages are 12, 13, and 14. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

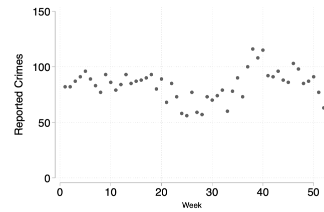
Figure A.14: Reported Crimes by Week, by Age of Offender and Age of Victim, Offenders Aged 15-17



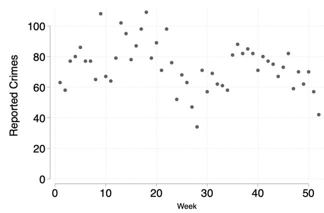
(a) O:15, V:12



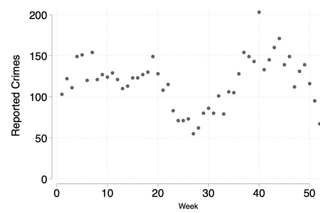
(b) O:16, V:13



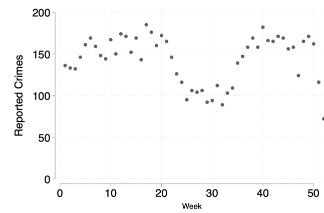
(c) O:17, V:14



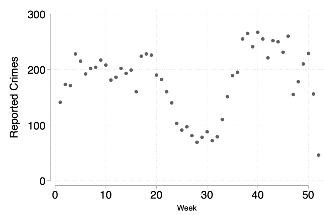
(d) O:15, V:13



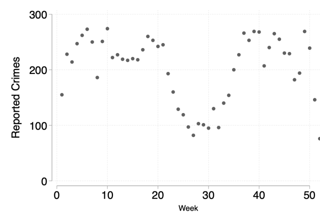
(e) O:16, V:14



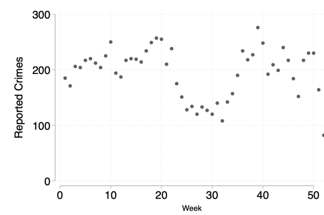
(f) O:17, V:15



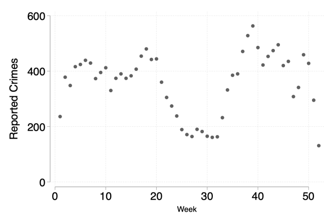
(g) O:15, V:14



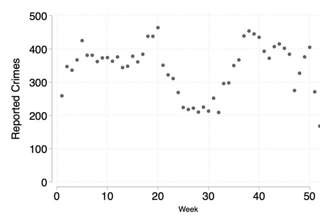
(h) O:16, V:15



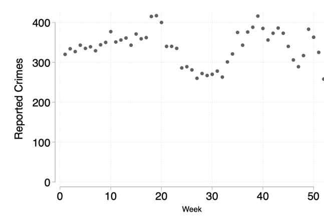
(i) O:17, V:16



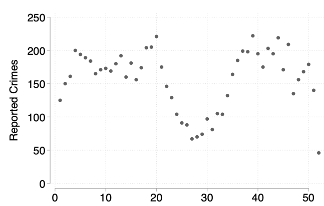
(j) O:15, V:15



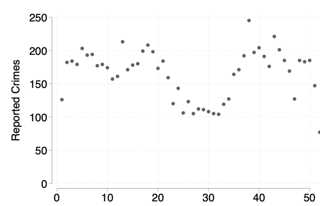
(k) O:16, V:16



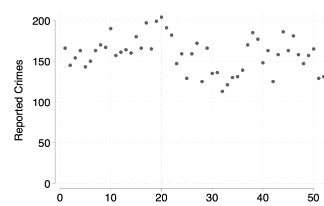
(l) O:17, V:17



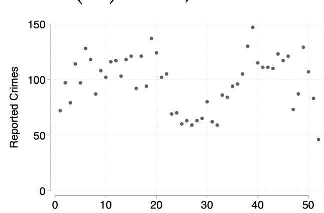
(m) O:15, V:16



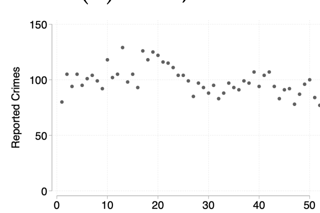
(n) O:16, V:17



(o) O:17, V:18



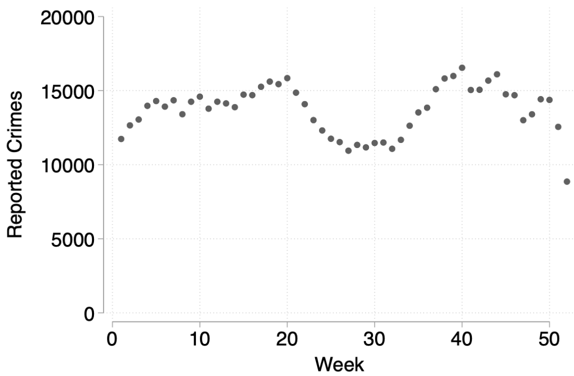
(p) O:15, V:17



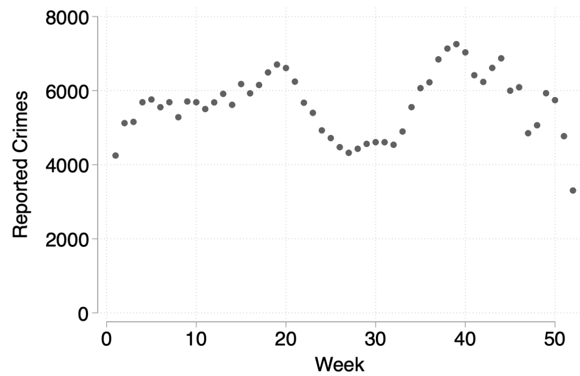
(q) O:16, V:18

Notes: These figures show the total number of reported crimes by week for age of first offender (O) and age of first victim (V) pairs. For example, “O: 15, V: 14” means that the offender was 15 and the victim was 14. Offender ages are 15, 16, and 17. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

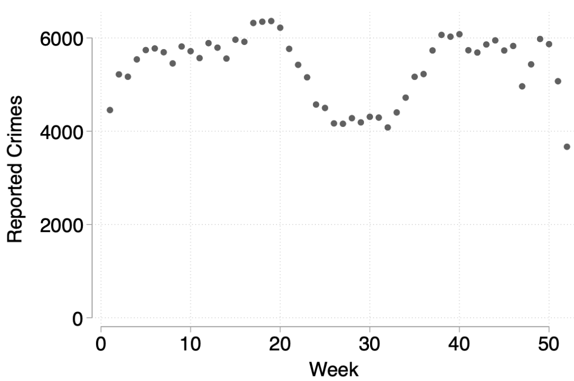
Figure A.15: Reported Crimes by Week, Age 10-17, by Male/Female and Offender/Victim



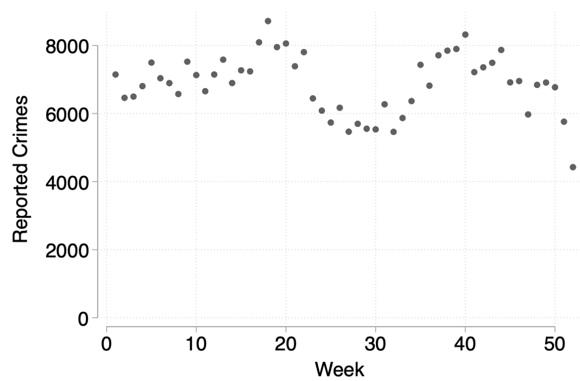
(a) Male, Offender



(b) Male, Victim



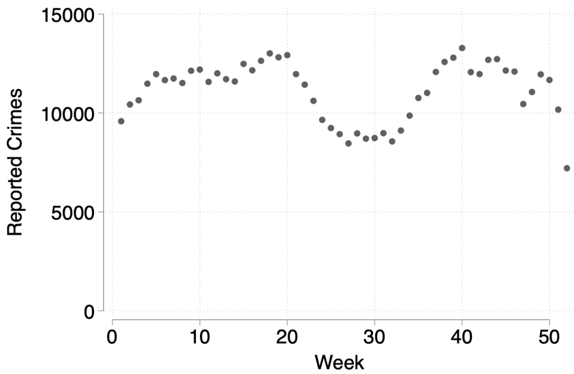
(c) Female, Offender



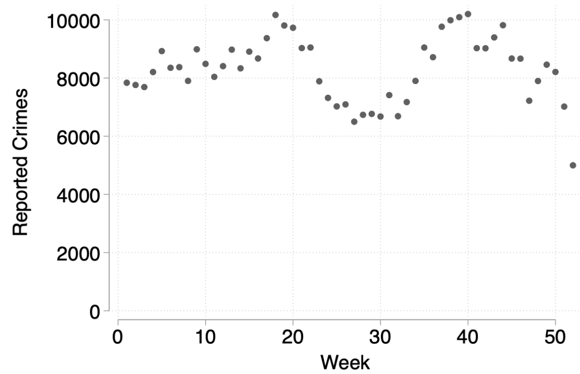
(d) Female, Victim

Notes: These figures show the total number of reported crimes by week. Each panel corresponds to a male/female -by- offender/victim pair. For offenders (Panels A and C), offender age is limited to 10-17. For victims (Panels B and D), victim age is limited to 10-17. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

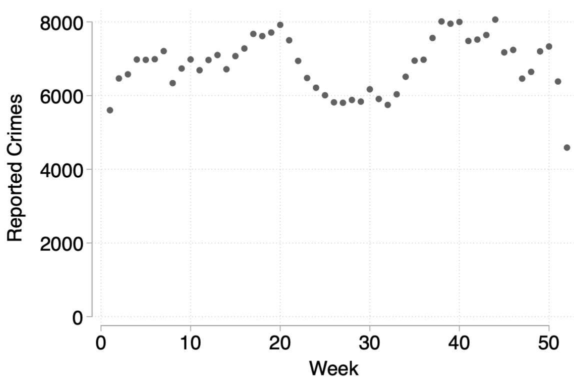
Figure A.16: Reported Crimes by Week, Age 10-17, by Black/White/Other Race and Offender/Victim



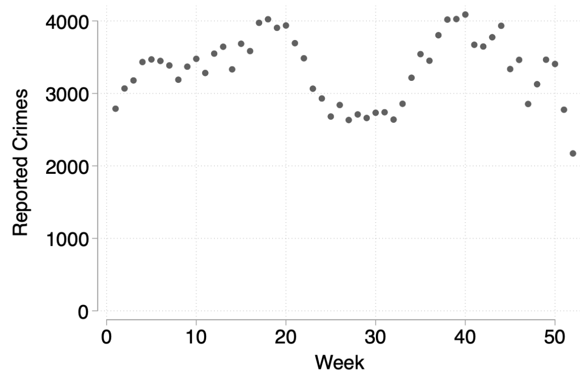
(a) White, Offender



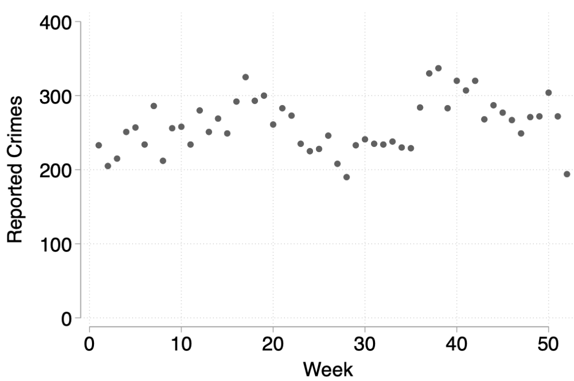
(b) White, Victim



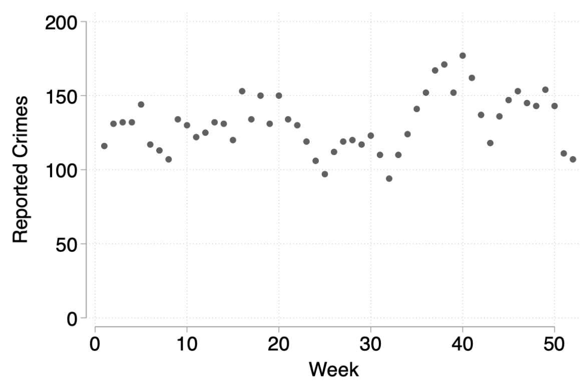
(c) Black, Offender



(d) Black, Victim



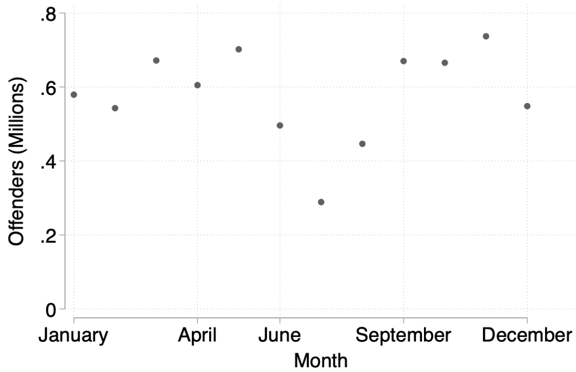
(e) Other Race, Offender



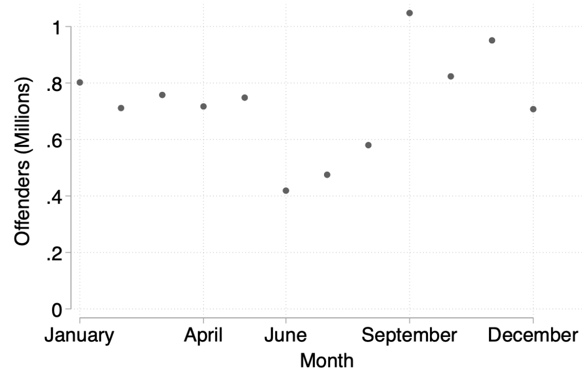
(f) Other Race, Victim

Notes: These figures show the total number of reported crimes by week. Each panel corresponds to a Black/White/Other Race+offender/victim pair. Other Race includes American Indian, Alaskan Native, Asian, and Pacific Islander, and Native Hawaiian. For offenders (Panels A, C, and E), offender age is limited to 10-17. For victims (Panels B, D, and F), victim age is limited to 10-17. The data source is 2017-2019 NIBRS. Week 1 starts on January 1.

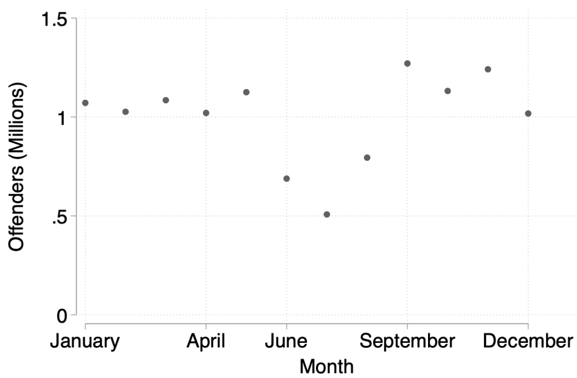
Figure A.17: Reported Crimes by Month, NCVS, Victims Age 10-17, by Type of Offense



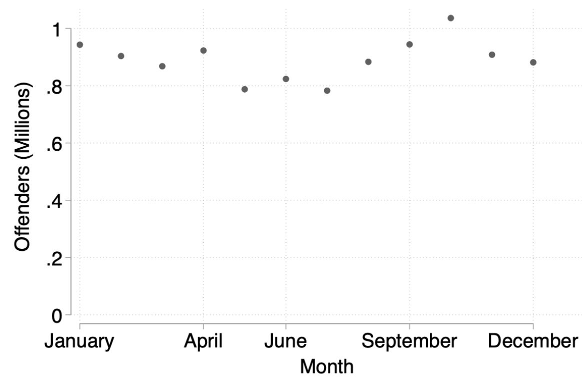
(a) Simple Assault



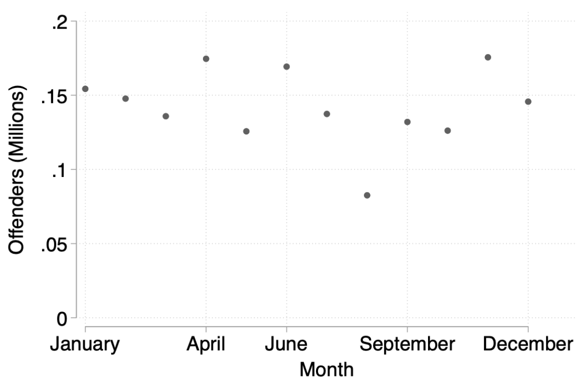
(b) Verbal Threat



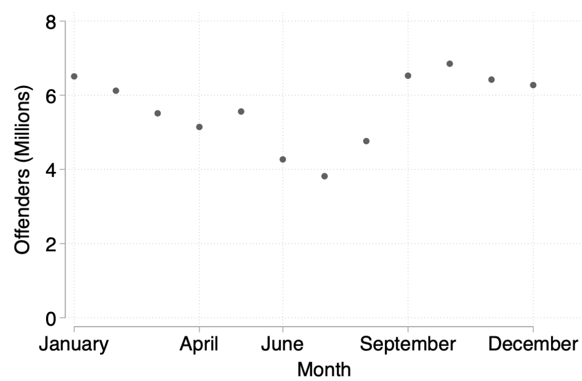
(c) Assault, No Weapon, No Injury



(d) Other Violent



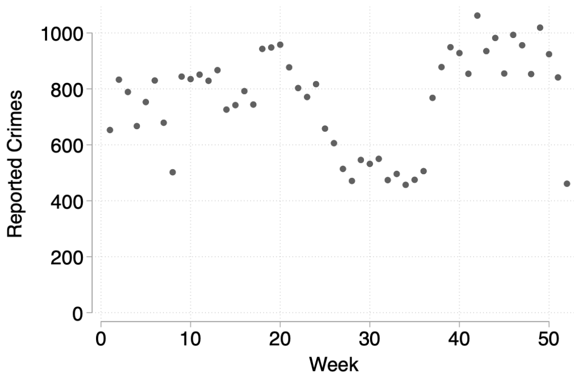
(e) Sexual Assault



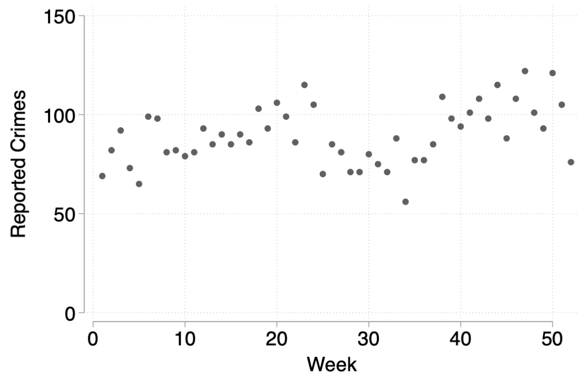
(f) Theft

Notes: These figures show the total number (after weighting using person weights) of reported victimizations by week for victims aged 10-17. Each panel corresponds to a different type of victimization. Panel a) shows simple assaults; Panel b) shows verbal threats; Panel c) shows assaults involving no weapon and no injury; Panel d) shows other violent crimes; Panel e) sexual assaults; and Panel f) shows thefts. The data source is 1992-2019 National Crime Victimization Survey. Week 1 starts on January 1.

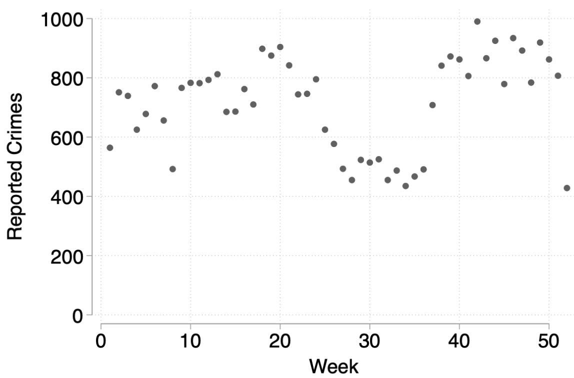
Figure A.18: Reported Crimes by Week, Age 17 and Under, New York City, < .1 Mile to Nearest School, by Public and Private, and by Schools with Few vs. Many Reported Crimes



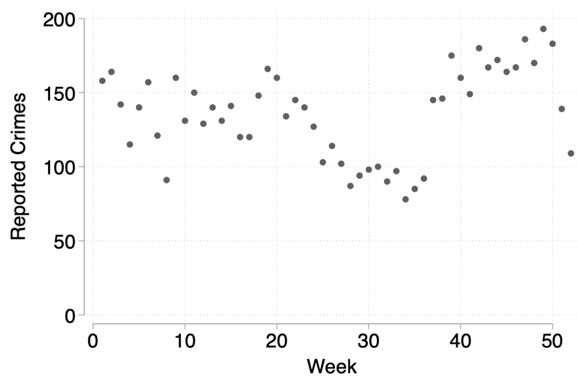
(a) Public Schools



(b) Private Schools



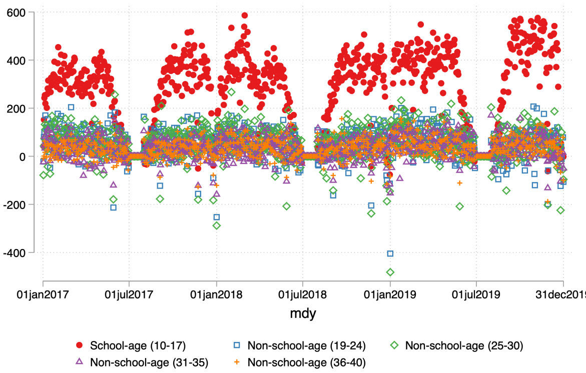
(c) Schools with <100 Total Reported Crimes



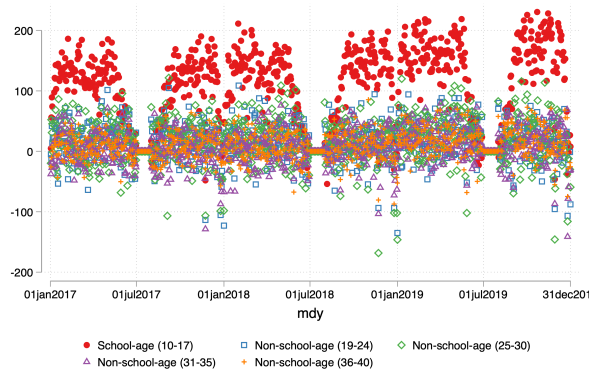
(d) Schools with ≥100 Total Reported Crimes

Notes: These figures show the total number of reported crimes by week for ages 17 and younger using New York City data. Each observation is matched, using latitudes and longitudes, to the nearest school geocode using data from NCES. All Panels are limited to observations occurring within 0.1 miles of the nearest school. Panel A is for public schools; Panel B is for private schools; Panel C is for schools with fewer than 100 observations over the entire sample; and Panel D is for schools with at least 100 observations over the entire sample. Observations are limited to those occurring between 2006-2019. Week 1 starts on January 1.

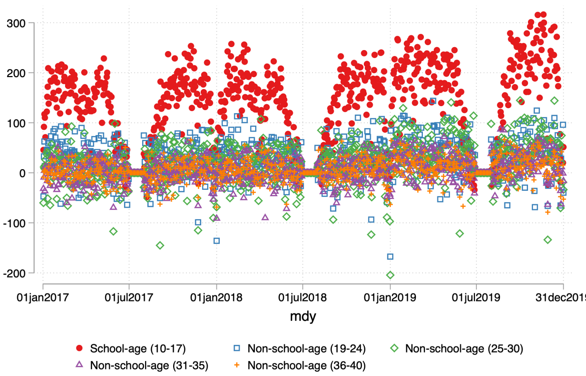
Figure A.19: Weekday Seasonally-adjusted Reported Crime, by Sex and Race



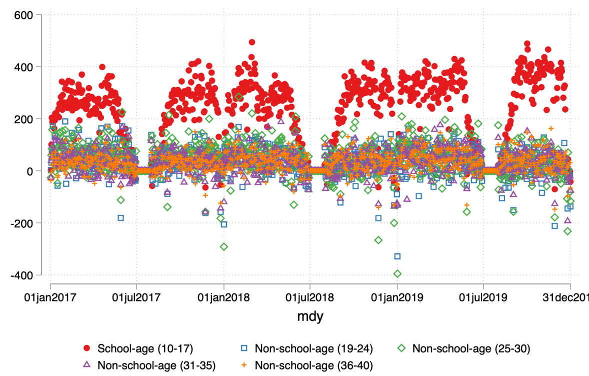
(a) Male Offenders



(b) Female Offenders



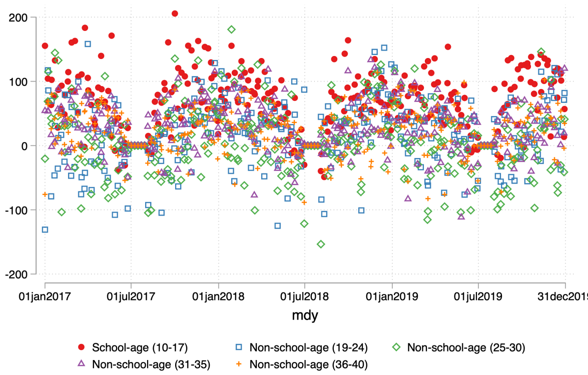
(c) Black Offenders



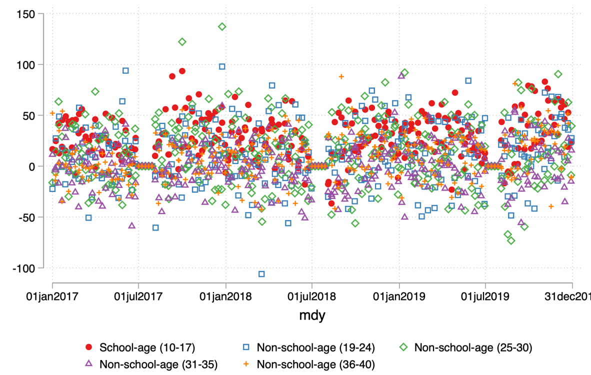
(d) White Offenders

Notes: Seasonal adjustment calculated by regressing daily counts of reported crime in the 2017–2019 NIBRS panel (by calendar day by- hour of the day -by- age of offender) on calendar day -by- age fixed effects and day-of-the-week -by- hour-of-the-day -by- age fixed effects for crime reported outside of the school calendar and school hours (in July or between 8pm and 6am). Subtracting these fixed effects from reported cell-level crime and summing the resulting residual by calendar day produces the (daily) residuals presented above (pegged to July 1st). Panel A is for male offenders, Panel B is for female offenders, Panel C is for Black offenders, and Panel D is for White offenders. Results are limited to weekdays.

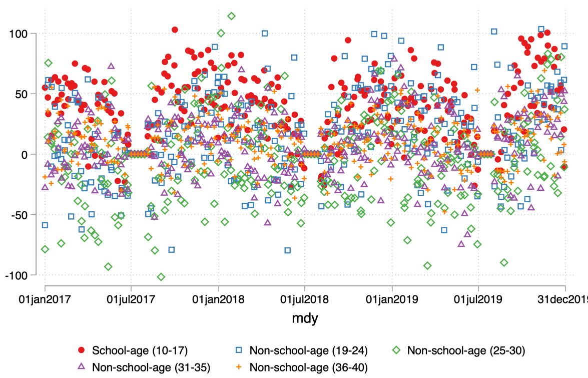
Figure A.20: Weekend Seasonally-adjusted Reported Crime, by Sex and Race



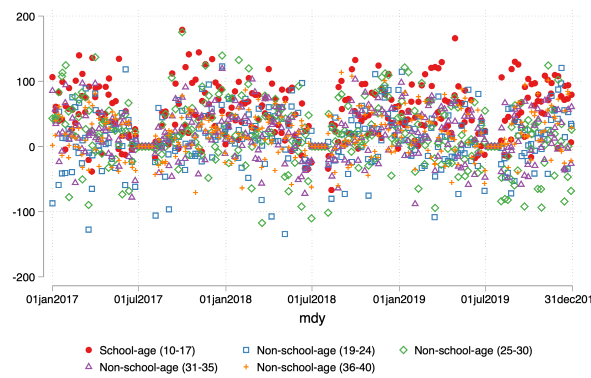
(a) Male Offenders



(b) Female Offenders



(c) Black Offenders



(d) White Offenders

Notes: Seasonal adjustment calculated by regressing daily counts of reported crime in the 2017–2019 NIBRS panel (by calendar day by- hour of the day -by- age of offender) on calendar day -by- age fixed effects and day-of-the-week -by- hour-of-the-day -by- age fixed effects for crime reported outside of the school calendar and school hours (in July or between 8pm and 6am). Subtracting these fixed effects from reported cell-level crime and summing the resulting residual by calendar day produces the (daily) residuals presented above (pegged to July 1st). Panel A is for male offenders, Panel B is for female offenders, Panel C is for Black offenders, and Panel D is for White offenders. Results are limited to weekends.

B Appendix Tables

Table B.1: RDD Estimates, All Incidents, Uniform Kernel

Panel A: School Start					
BW:	3	4	6	8	10
10–17	585.67	695.00	815.40***	934.39***	1099.73***
	(437.82)	(434.24)	(271.79)	(210.10)	(173.37)
Summer Mean:	2,051.00	2,093.00	2,109.33	2,128.00	2,160.40
19–24	-261.33***	-312.50***	-123.53	11.75	-84.47
	(91.28)	(108.67)	(92.74)	(98.27)	(67.73)
Summer Mean:	5,026.67	5,015.25	5,039.17	5,108.50	5,103.40
25–30	204.00	218.00	167.93	95.07	116.33
	(236.84)	(211.57)	(172.88)	(128.71)	(112.98)
Summer Mean:	5,819.67	5,845.75	5,847.83	5,868.62	5,893.60
Panel B: School End					
BW:	3	4	6	8	10
10–17	-660.00***	-686.50***	-892.73***	-990.21***	-1008.67***
	(229.31)	(216.97)	(159.11)	(111.38)	(95.97)
School Mean:	3,226.00	3,265.50	3,202.33	3,154.12	3,108.90
19–24	-260.33	-188.00	-195.20	-142.64	-100.27
	(181.76)	(133.53)	(135.01)	(111.06)	(107.57)
School Mean:	5,282.33	5,298.75	5,245.17	5,217.00	5,204.40
25–30	250.00	153.00	64.87	52.93	43.47
	(257.38)	(163.29)	(119.84)	(103.11)	(90.40)
School Mean:	5,967.00	5,948.50	5,910.83	5,891.50	5,845.00

This tables shows estimates of all incidents from the regression discontinuity design, with a uniform kernel, obtained using rdrobust (Calonico et al., 2017). Standard errors are displayed in parentheses. Each observation represents the number of total incidents (Panel A) and arrests (Panel B) in a given week. Each estimate comes from a separate regression, with results presented separately by age bin and bandwidth (number of weeks). Panel A (B) corresponds to the school start (end) cutoff. The mean of the last weeks of the summer is presented for Panel A, and the mean of the last weeks of the school year is presented in Panel B.

Table B.2: RDD Estimates, All Incidents, Triangular Kernel

Panel A: School Start					
BW:	3	4	6	8	10
10–17	253.00	502.50	701.40*	784.71**	900.17***
	(549.56)	(473.52)	(397.73)	(300.61)	(244.83)
Summer Mean:	2,051.00	2,093.00	2,109.33	2,128.00	2,160.40
19–24	-138.00	-230.50**	-257.00***	-123.64	-73.17
	(116.29)	(92.69)	(91.43)	(85.37)	(76.74)
Summer Mean:	5,026.67	5,015.25	5,039.17	5,108.50	5,103.40
25–30	-118.00	123.50	161.20	152.75	121.73
	(381.22)	(235.64)	(181.90)	(157.02)	(135.21)
Summer Mean:	5,819.67	5,845.75	5,847.83	5,868.62	5,893.60

Panel B: School End					
BW:	3	4	6	8	10
10–17	-524.00*	-626.00**	-747.20***	-872.57***	-935.63***
	(306.28)	(250.16)	(206.77)	(155.67)	(126.25)
School Mean:	3,226.00	3,265.50	3,202.33	3,154.12	3,108.90
19–24	67.00	-178.50	-221.20*	-186.89	-135.37
	(232.57)	(137.36)	(118.23)	(113.53)	(107.98)
School Mean:	5,282.33	5,298.75	5,245.17	5,217.00	5,204.40
25–30	284.00	258.50	130.80	78.50	63.70
	(334.22)	(281.08)	(158.14)	(123.97)	(107.28)
School Mean:	5,967.00	5,948.50	5,910.83	5,891.50	5,845.00

This tables shows estimates of all incidents from the regression discontinuity design, with a triangular kernel, obtained using rdrobust (Calonico et al., 2017). Standard errors are displayed in parentheses. Each observation represents the number of total incidents (Panel A) and arrests (Panel B) in a given week. Each estimate comes from a separate regression, with results presented separately by age bin and bandwidth (number of weeks). Panel A (B) corresponds to the school start (end) cutoff. The mean of the last weeks of the summer is presented for Panel A, and the mean of the last weeks of the school year is presented in Panel B; we report the mean corresponding to the bandwidth used, even if the effective number of observations used in estimation differs.

Table B.3: RDD Estimates, Arrests, Uniform Kernel

Panel A: School Start					
BW:	3	4	6	8	10
10–17	41.00	64.50	103.87	124.96*	161.60***
	(129.95)	(116.78)	(79.10)	(63.77)	(52.01)
Summer Mean:	567.00	572.50	573.17	573.88	579.70
19–24	105.00	24.50	57.13	40.18	8.73
	(128.91)	(89.78)	(80.95)	(62.76)	(50.16)
Summer Mean:	1,951.00	1,959.50	1,991.17	2,000.38	2,006.80
25–30	212.33	235.50	112.80	31.29	24.33
	(162.58)	(179.24)	(121.19)	(90.08)	(79.29)
Summer Mean:	2,379.00	2,401.75	2,388.00	2,389.62	2,394.60
Panel B: School End					
BW:	3	4	6	8	10
10–17	-83.33	-60.50	-120.87**	-156.29***	-168.07***
	(58.80)	(71.13)	(45.19)	(35.32)	(32.83)
School Mean:	744.33	763.50	746.33	735.12	725.70
19–24	2.00	-50.50	-48.20	-30.64	-0.00
	(95.93)	(80.55)	(73.60)	(52.37)	(52.30)
School Mean:	1,962.00	1,965.00	1,954.33	1,954.00	1,948.30
25–30	158.67	134.00	38.07	-18.79	8.73
	(127.30)	(102.68)	(72.08)	(61.67)	(50.49)
School Mean:	2,284.67	2,294.75	2,291.17	2,279.50	2,278.50

This tables shows estimates of arrests from the regression discontinuity design, with a uniform kernel, obtained using rdrobust (Calonico et al., 2017). Standard errors are displayed in parentheses. Each observation represents the number of total incidents (Panel A) and arrests (Panel B) in a given week. Each estimate comes from a separate regression, with results presented separately by age bin and bandwidth (number of weeks). Panel A (B) corresponds to the school start (end) cutoff. The mean of the last weeks of the summer is presented for Panel A, and the mean of the last weeks of the school year is presented in Panel B.

Table B.4: RDD Estimates, Arrests, Triangular Kernel

Panel A: School Start					
BW:	3	4	6	8	10
10–17	-175.00	-13.00	61.60	82.71	115.90
	(242.82)	(141.90)	(108.45)	(84.99)	(71.00)
Summer Mean:	567.00	572.50	573.17	573.88	579.70
19–24	-45.00	67.50	15.40	38.54	34.07
	(93.43)	(126.40)	(85.86)	(72.28)	(60.89)
Summer Mean:	1,951.00	1,959.50	1,991.17	2,000.38	2,006.80
25–30	185.00*	205.50	179.40	111.54	61.13
	(110.68)	(168.37)	(155.64)	(119.80)	(98.87)
Summer Mean:	2,379.00	2,401.75	2,388.00	2,389.62	2,394.60
Panel B: School End					
BW:	3	4	6	8	10
10–17	-12.00	-65.50	-75.20	-116.71**	-141.77***
	(94.34)	(64.00)	(65.41)	(47.38)	(38.59)
School Mean:	744.33	763.50	746.33	735.12	725.70
19–24	186.00	48.00	-36.60	-39.43	-19.33
	(190.22)	(95.44)	(68.08)	(62.50)	(56.13)
School Mean:	1,962.00	1,965.00	1,954.33	1,954.00	1,948.30
25–30	126.00	150.50	90.00	49.14	18.10
	(140.65)	(135.68)	(98.57)	(75.94)	(64.00)
School Mean:	2,284.67	2,294.75	2,291.17	2,279.50	2,278.50

This tables shows estimates of arrests from the regression discontinuity design, with a triangular kernel, obtained using rdrobust (Calonico et al., 2017). Standard errors are displayed in parentheses. Each observation represents the number of total incidents (Panel A) and arrests (Panel B) in a given week. Each estimate comes from a separate regression, with results presented separately by age bin and bandwidth (number of weeks). Panel A (B) corresponds to the school start (end) cutoff. The mean of the last weeks of the summer is presented for Panel A, and the mean of the last weeks of the school year is presented in Panel B; we report the mean corresponding to the bandwidth used, even if the effective number of observations used in estimation differs.

Table B.5: Excess Crime, Control Group 19–24 Year Olds

Panel A: All Incidents	Excess Crime	% Increase
All	0.31	0.45
White	0.32	0.47
Black	0.30	0.43
Male	0.31	0.45
Female	0.32	0.46

Panel B: Arrests	Excess Crime	% Increase
All	0.28	0.39
White	0.28	0.39
Black	0.28	0.38
Male	0.28	0.39
Female	0.27	0.38

This table reports how much 'excess crime' there is for 10–17 year olds, for all incidents in Panel A, and for arrests in Panel B. Each line corresponds to a different sample. To obtain the estimates, we perform a seasonal adjustment calculated by regressing daily counts of reported crime in the 2017–2019 NIBRS panel (by calendar day by- hour of the day -by- age of offender) on calendar day -by- age fixed effects and day-of-the-week -by- hour-of-the-day -by- age fixed effects for crime reported outside of the school calendar and school hours (in July or between 8pm and 6am). Subtracting these fixed effects from reported cell-level crime and summing the resulting residual by calendar day for produces the (daily) residuals for each age group. We then compute the average residual for each age group. We subtract the mean from the control group, 19–year olds, after scaling it to account for differences in sample size between it and the 10–17 year old group. We multiply this object by 3*365 and to account for the number of days in the three years (call this Z), which we divide by the total number of crimes for the 10–17 year old group (call this N) to obtain the number in the 'Excess Crime' column. The '% Increase' column is obtained by dividing Z by (N minus Z).

Table B.6: County Excess Crime Predictors, Unweighted

Poverty %	0.10	0.17	0.17	0.17
	(0.12)	(0.13)	(0.14)	(0.11)
Unemployment Rate	0.02	0.02	0.02	0.02
	(0.07)	(0.10)	(0.09)	(0.08)
Median HH Income	0.03	0.06	0.06	0.06
	(0.12)	(0.13)	(0.08)	(0.05)
Population	-0.03	-0.05	-0.05	-0.05
	(0.07)	(0.07)	(0.05)	(0.05)
% Less than High School	-0.08	-0.10	-0.10	-0.10
	(0.08)	(0.10)	(0.18)	(0.16)
Rural [0,1]	0.13	0.08	0.08	0.08
	(0.16)	(0.16)	(0.13)	(0.11)
Expenditures per Student	0.11*	0.23**	0.23	0.23
	(0.07)	(0.10)	(0.24)	(0.23)
Students per Grade	-0.03	-0.05	-0.05	-0.05
	(0.09)	(0.12)	(0.08)	(0.08)
% White and Asian	-0.18**	-0.19*	-0.19	-0.19
	(0.07)	(0.10)	(0.20)	(0.21)
Mixed HS Jr High	-0.15**	-0.15*	-0.15	-0.15*
	(0.07)	(0.08)	(0.09)	(0.08)
Any SRO or Security Guard	0.12*	0.14*	0.14	0.14
	(0.07)	(0.08)	(0.10)	(0.09)
Y Mean	0.33	0.33	0.33	0.33
N	1,824	1,820	1,820	1,820
R-squared	0.01	0.05	0.05	0.05
State FEs	No	Yes	Yes	Yes
SEs			Robust	State Cluster

This table reports a regression of county-level excess crime on a number of variables. It does not use county weights. Unless indicated by '[0,1]', each variable is standardized (among the 1,824 counties) to have mean 0 and standard deviation 1. The first column does not include fixed effects or adjustments for standard errors. Columns 2–4 include state fixed effects. Column 3 uses robust standard errors. Column 4 clusters the standard errors at the state level. * $<.1$; ** $<.05$; *** $<.01$.

Table B.7: County Excess Crime Predictors, Weight by Grade 5-12 Enrollment

Poverty %	-0.01	0.03	0.03	0.03
	(0.05)	(0.05)	(0.05)	(0.04)
Unemployment Rate	-0.00	-0.02	-0.02	-0.02
	(0.03)	(0.04)	(0.04)	(0.05)
Median HH Income	-0.05	-0.05	-0.05	-0.05
	(0.03)	(0.03)	(0.03)	(0.03)
Population	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.00)
% Less than High School	-0.05	-0.01	-0.01	-0.01
	(0.04)	(0.04)	(0.04)	(0.05)
Rural [0,1]	0.06	0.06	0.06	0.06
	(0.08)	(0.08)	(0.08)	(0.06)
Expenditures per Student	-0.00	0.02	0.02	0.02
	(0.02)	(0.05)	(0.05)	(0.05)
Students per Grade	0.08	0.10*	0.10*	0.10
	(0.05)	(0.06)	(0.06)	(0.08)
% White and Asian	-0.05	0.00	0.00	0.00
	(0.04)	(0.04)	(0.04)	(0.04)
Mixed HS Jr High	-0.10*	-0.07	-0.07	-0.07
	(0.06)	(0.07)	(0.07)	(0.07)
Any SRO or Security Guard	-0.11*	-0.04	-0.04	-0.04
	(0.05)	(0.04)	(0.04)	(0.07)
Y Mean	0.33	0.33	0.33	0.33
N	1,824	1,820	1,820	1,820
R-squared	0.01	0.07	0.07	0.07
State FEs	No	Yes	Yes	Yes
SEs			Robust	State Cluster

This table reports a regression of county-level excess crime on a number of variables. It weights counties by grade 5–12 enrollment. Unless indicated by '[0,1]', each variable is standardized (among the 1,824 counties) to have mean 0 and standard deviation 1. The first column does not include fixed effects or adjustments for standard errors. Columns 2–4 include state fixed effects. Column 3 uses robust standard errors. Column 4 clusters the standard errors at the state level. * <.1; ** <.05; *** <.01.

Table B.8: County Excess Crime Predictors, Weight by Number of Reported Crimes among 10-40 Year Olds

Poverty %	-0.02	0.01	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Unemployment Rate	-0.00	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Median HH Income	-0.02	-0.02	-0.02	-0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
Population	-0.00*	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
% Less than High School	-0.02	-0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Rural [0,1]	0.02	0.01	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Expenditures per Student	-0.00	-0.01	-0.01	-0.01
	(0.01)	(0.02)	(0.02)	(0.02)
Students per Grade	0.05***	0.04**	0.04**	0.04**
	(0.01)	(0.02)	(0.02)	(0.02)
% White and Asian	-0.02	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Mixed HS Jr High	-0.02	-0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Any SRO or Security Guard	-0.03*	-0.01	-0.01	-0.01
	(0.02)	(0.01)	(0.01)	(0.02)
Y Mean	0.33	0.33	0.33	0.33
N	1,824	1,820	1,820	1,820
R-squared	0.01	0.14	0.14	0.14
State FEs	No	Yes	Yes	Yes
SEs			Robust	State Cluster

This table reports a regression of county-level excess crime on a number of variables. It weights counties by the number of reported crimes in the data at the county level among 10–40 year olds, after excluding observations missing date, time, age, or occurring between midnight and 1am. Unless indicated by '[0,1]', each variable is standardized (among the 1,824 counties) to have mean 0 and standard deviation 1. The first column does not include fixed effects or adjustments for standard errors. Columns 2–4 include state fixed effects. Column 3 uses robust standard errors. Column 4 clusters the standard errors at the state level. * <.1; ** <.05; *** <.01.

C Data Appendix

C.1 NIBRS

The NIBRS data consists of several files, including the *Administrative Segment*, *Offense*, *Offender*, *Victim*, *Arrestee*, and *Batch Header* files. We combine these files together. NIBRS also contains *Group B* arrest files, which we process separately as they are stand-alone records. We draw from Kaplan (2021b) for some of the information below. Below, we outline the steps involved in cleaning this data:

Administrative Segment file: We begin with this file, which contains baseline information for each incident. Within year, the data are organized at the ORI (police department) and incident number level, which should be a unique identifier. However, for the 2017 and 2018 data, less than 1% of ORI-incident number pairs (before sample selection) are not unique at this level; most of these cases are, however, unique at the ORI-incident number-date level; we assume that these are in fact unique incidents and so process the data at this level. After dropping the handful of cases missing incident number, we drop the small percentage of incidents that are not unique at the ORI-incident number-date level because there is not enough information in the other files with which to correctly match. In 2017, the number of observations that are not unique and that we therefore drop is only 0.11% (about one-tenth of one percent). For 2018, this number is 0.12%. While these observations are not uniformly distributed across date (they are more likely to be from July through September), the fact that they make up just a sliver of the overall data make it very unlikely that not considering these observations will meaningfully affect results. From this file we obtain the date that we assign to the incident; in most cases (about 90%, depending on the sample), this is the date the incident occurred, but in the remaining cases it is the date the incident was reported. We also note that the first day of the month occurs more than would be expected (Kaplan, 2021b); we do not exclude these cases. In addition, the midnight–12:59am and 12:00pm-12:59pm hours occur more than would be expected (Kaplan, 2021b); we generally do not exclude these cases, but do in the excess crime calculation.

Batch Header file: This file provides information about the ORI, including the number of months it reported data.

Offense file: This file contains data on offenses, including UCR offense codes and locations. Sometimes, there are multiple offenses within a single incident. In these cases, we keep the observation that we first observe in the data, with the exception that we prioritize observations with (in order) a location of "school - elementary/secondary", "school/college", and "school - college/university." When we consider the location of the crime, we define school as being "school - elementary/secondary" and "school/college", but not "school - college/university."

Offender file: The *Offender* file contains data on offenders, such as their age, sex, and race. Not all observations have complete information, similar to some variables in the files listed below. The data in this file are organized at the ORI-incident number-date-offender sequence level. We construct the number of offenders per incident. We drop if offender sequence level is unknown; offender age is never observed in these instances. We additionally drop observations where the offender is less than 10 years old. We then keep the offender with the youngest age, breaking ties by first-occurring offender sequence number.

Victim file: The *Victim* file provides data on the victims, including the offense code type, victim age, sex, race, and the relationship between the victim and the offender(s). The data in this file are organized at the ORI-incident number-date-victim sequence level. We keep the victim with the youngest age, breaking ties by first-occurring victim sequence number. We also consider only the first-listed offense code. We consider the relationship between the victim and the offender from the offender file.

Arrestee file: This file contains information about the arrestees, such as the date of arrest (which may differ from the incident date), age, sex, and race. It also includes the type of arrest, focusing on "on-view arrest" (custody without a warrant or previous incident report) and "taken into custody" (based on a warrant or previous incident report). In our main results, we exclude "summoned/cited" cases (not taken into custody), but we present results for this category in the Appendix. The data in this file are organized at the ORI-incident number-date-arrestee sequence level. We additionally drop observations where the arrestee is less than 10 years old. We then keep the arrestee with the youngest age, breaking ties by first-occurring arrestee sequence number. Sometimes arrests clear multiple incidents; in these cases, we consider only the incident coded as the first, or 'count arrestee' incident. When we consider arrests, we use the incident date as opposed to the arrest date; the former occurs on the same day or on a day before the latter. Note also that, for a given observation, the person we consider for the 'all incidents' sample is not always the same person as the arrestee for this incident. For instance, if an incident had two offenders, a 15 year old and a 17 year old, where only the 17 year old was arrested. We would consider the 15 year old in the 'all incidents' sample as they are the youngest (see above), but the 17 year old in the 'arrest' sample because they are the only person from this incident who was arrested.

After combining the 2017, 2018, and 2019 data together for each file type, we merge the *Offense*, *Offender*, *Victim*, *Arrestee*, and *Batch Header* segments onto the *Administrative Segment*. We additionally merge name of district and geographical data, including state code, from the *Law Enforcement Agency Identifiers Crosswalk, United States, 2012* (Bureau of Justice Statistics, 2018). We rely on the county FIPS codes from the NIBRS data for analysis that involves a county (in the *Batch Header* file, a relatively small percentage of ORIs are associated with multiple county FIPS codes indicating that they may cross county lines). Offense code is included in both the *Offender* and *Victim* files; in some cases, these differ as there can be multiple offense codes for an incident. We rely on the offense code associated with the victim.

Group B Arrests: Group B arrests are not associated with an incident, and so we consider them separately in the analysis. We use the data of arrest as the date of the incident is not provided. Group B categories include bad checks, curfew/loitering/vagrancy violations, disorderly conduct, driving under the influence, drunkenness, family offenses nonviolent, liquor law violations, peeping tom, runaway, and trespass of real property. They also include all other offenses, which before sample selections make up more than half of all Group B observations. These can include attempted crimes that are not completed.

C.2 RDD Sample

Because the school calendar data does not have geographic identifies besides state, we merge geographic information from NCES (National Center for Education Statistics, 2020) onto the school calendar data using fuzzy matching based on state and district name. We first clean both files. In the school calendar file, we keep dates from the 2019–2020 school year, and rename one district to resolve an issue that would arise later in the process. In the NCES file, some district names are duplicated across counties within state; we drop these instances (less than 1% of observations; we drop all but one before matching and all observations matched to the remaining one after matching). For both files, we convert the district name to lower case; remove strings such as district, school, schools, dist., ccsd, elementary, elem, and regional; and replace numbers like 01 to 1. If at this point, there are new duplicates that the state-district name level, we revert back to the original (lower case) name. The fuzzy matching yields 11,174 matches of 11,245 total, including 8,601 perfect matches (76%). We drop unmatched observations, observations in which a school calendar school district matches to multiple NCES school districts that are in different counties, NCES school districts that are matched to multiple school calendar school districts, and observations matched to NCES school districts that had a name was duplicated across counties, keeping only one observation per state-district name. These restrictions collectively drop 152 observations, leaving us with 11,022 school district observations.

We then match the school calendar information onto the NIBRS incident-level file, doing so at the county level. To improve accuracy, we restrict our attention to the counties for which all observed start (end) dates for all school districts in the sample fall on the same week. One complication is that about 4% of the ORI's (police departments) contain more than one FIPS code (span multiple counties). In these cases, we require that all school districts in the sample across all FIPS codes have the same start (end) week. The prior paragraph described several cases in which we removed districts from the sample; doing so improves measurement of district start (end) week, but also removed potentially valid cases; to the extent to which the removed districts have differing start (end) weeks, we will keep incidents that have multiple start (end) weeks; however, this is unlikely to be much of a problem given that the number of districts we removed were relatively few. Another thing to note is that while the school calendar data contains a large number of school districts, it does not contain all of them (and some individual schools are in counties different than the county of their district); this may also lead to measurement error.

C.3 National Crime Victimization Survey

We obtain the National Crime Victimization Survey (NCVS) data from ICPSR (Bureau of Justice Statistics., 2021). We weight our results by the person weight.

We consider crimes with incident location "School building" and "School property" as having occurred at school and all other locations as not.

We create six categories of types of offense, and classify them as following: 1) Simple assault: simple assault completed with injury; 2) Verbal threat: Verbal threat of assault; 3) Assault, no Weapon, no Injury: assault without weapon without injury; 4) Other violent: completed robbery with injury from serious assault, completed robbery with injury from minor assault, completed robbery without injury from minor assault, attempted robbery with injury from serious assault, attempted robbery with injury from minor assault, attempted robbery without injury, completed aggravated assault with injury, attempted aggravated assault with weapon, threatened assault with weapon; 5) Sexual assault: completed rape, attempted rape, sexual attack with serious assault, sexual attack with minor assault, sexual assault without injury, unwanted sexual contact without force; 6) Theft: completed purse snatching, attempted purse snatching, pocket picking (completed only), completed burglary forcible entry, completed burglary unlawful entry without force, attempted forcible entry, completed motor vehicle theft, attempted motor vehicle theft, completed theft less than $10, completed theft $10 to $49, completed theft $50 to $249, completed theft $250 or greater, completed theft value NA, attempted theft.

C.4 New York City Data

We use the 'NYPD Complaint Data Historic' available from the New York Open Data citepnypd.[47] Only one offender is listed. We do not observe exact age of offender and so rely on the under 18 year old category. This includes people under 10 years old who would not be included in our main NIBRS sample. We use crimes reported to have occurred during the years 2006 to 2019. Geocodes are approximate (and the documentation says "Any attempt to match the approximate location of the incident to an exact address or link to other datasets is not recommended"), and in some cases to further prevent the indentify of the vistim—such as for rape victims—are geocoded at the police station within the precint where the crime occurred. Using geocoordinates for crimes and schools, we link incidents to the nearest school (public or private) as listed in the National Center for Education Statistics (NCES) Education Demographic and Geographic Estimates Public School and Private School files from 2019–2020 (National Center for Education Statistics, 2020).

[47] See https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i for notes about the data.

C.5 County Predictors Data

We first detail the sources of and procedures used to construct our variables. We then discuss the county-level excess crime and sample construction.

Poverty %

We use the 2018 percentage of all people in poverty variable from the 2018 USDA ERS poverty file (USDA Economic Research Service, 2019c).

Unemployment Rate

We use the 2018 unemployment rate from the 2018 USDA ERS unemployment file (USDA Economic Research Service, 2019d).

Median HH Income

We use the 2018 median household income variable from the 2018 USDA ERS unemployment file (USDA Economic Research Service, 2019d). *Population*

We use the 2018 population estimate from the ESDA ERS population file (USDA Economic Research Service, 2019b).

% Less than High School

We use the "Percent of adults with less than a high school diploma, 2014-18" variable from the USDA ERS file (USDA Economic Research Service, 2019a),

Rural

We use the USDA ERS Rural-Urban Continuum Codes (USDA Economic Research Service, 2020) to construct an indicator for the county being rural. Specifically, we consider the counties coded as 'nonmetro' to be rural, with the remaining ('urban') counties to not be rural.

Grade 5-12 Enrollment

We obtain grade 5-12 enrollment from the 2017 Common Core of Data (CCD) school enrollment file (Urban Institute, 2023); we use this variable as a weight in variable construction below. After restricting to grades 5–12, we aggregate enrollment to the county level. We do not consider the few cases where enrollment is not separately split by grade.

We start with the school file instead of the district file because in rare cases, schools are in a different county than their district. One caveat to this approach (which we also use to construct other variables involving enrollment) is that the enrollment reported in the district file does not match school-level enrollment aggregated to the district level, with the latter usually being higher in cases of disagreement; however, the correlation is close to 1. There are also relatively few cases of observations in the district file with no corresponding school files.

Educational Expenditures per Student

In the 2017 CCD data, spending is only available at the district level. We therefore construct educational spending per student at the district level. We do this by dividing the district "total current expenditures for elementary and secondary education" variable by district enrollment (as opposed to school enrollment aggregated to the district level). We then merge this onto the school data and aggregate to the county level, weighting by school grade 5–12 enrollment.

We also note that we investigated the "pp_total_norm_NERDS" variable from the National Education Resource Database on Schools (NERDS), which is a school-level spending variable made to be comparable across the United States (Edunomics Lab at Georgetown University, 2021). We elected to not focus on this variable as there are more missing observations at the county level compared to the CCD-derived variable described above. After aggregating to the county level (using the 2018 CCD data (Urban Institute, 2023) to obtain school-level county FIPS, and where this was not possible, district-level FIPS), we found that the correlation was 0.91 with the CCD version among the counties used in the analysis for which the NERDS variable was not missing; results are also qualitatively similar.

Number of Students per Grade

From the 2017 CCD school enrollment file, we restrict to grades 5–12 and aggregate enrollment and the number of grades (with at least one student) to the county level. We then divide the two.

Percentage of Students who are White or Asian

Using the 2017 CCD school enrollment file, we compute the percentage of students in grades 5–12 that are white or Asian; this category excludes two or more races. We drop those with unknown race. We aggregate to the county level, weighting by grade 5–12 enrollment.

Mixed HS Jr High

From the 2017 CCD school enrollment file, we restrict to grades 6–12, representing typical grades for middle school (6–8) and high school (9–12). We then construct a minimum grade variable, which is the lowest grade with at least one student, and a maximum grade variable, which is the highest grade with at least one student. At the school level, we then compute A) the total enrollment of schools containing high schoolers, defined as the total enrollment of schools with a maximum grade between 9 and 12, and B) the total enrollment of schools containing both middle schoolers and high schoolers, defined as the total enrollment of schools with both a minimum grade between 6 and 8 and a maximum grade between 9 and 12. We then divide B by A to produce a measure of mixed high school and junior high schools. Only schools with higher schoolers will have a value of this variable, and the value will either be 0, for high schools that don't contain at least one middle schooler, and 1, for high schools that do contain at least one middle schooler. We then aggregate this variable to the county level, weighting by grade 9–12 enrollment. The binary variable becomes continuous in this step. We assign the rare counties that do not have high schoolers in the school enrollment file with a value of 0.

School Resource Officers

We obtain data on school resource officers (SROs) from the 2017-18 US Department of Education Civil Rights Data Collection (CRDC) (U.S. Department of Education, 2018). We use the FTE security law enforcement officer (we refer to this as SRO, or school resource officer) and FTE security guard variables. At the school level, we construct an indicator for the school having at least one SRO or security guard (even if this person is not full time). We aggregate to the county level, weighting by school enrollment. The binary variables become continuous in this step.

We now describe additional data cleaning and matching variables to accomplish the above. We obtain the LEO and security guard variables from the CRDC. We do not consider schools that are a juvenile justice facility. While the CRDC data contains school-level enrollment, it does not contain school-grade level enrollment data, which is preferred so that we can limit to grade 5–12 enrollment. We therefore use the school-grade level data from the 2017 CCD (Urban Institute, 2023). We use a crosswalk to go between the CRDC and CCD datasets (American Institutes for Research, 2020). This matches the vast majority of schools, but not all as the coverage differs between datasets, where each dataset contains schools not covered by the other. When weighting by enrollment when aggregating to the county level, we primarily use the CCD enrollment data (and drop the observation if there is 0 CCD enrollment for the school) The exceptions are that we instead use the CRDC enrollment data (scaled by the number of grades between 5 through 12 that are represented in the school) when there is no match between the CRDC and CCD datasets and when CRDC schools are not unique using the CCD school code obtained from the crosswalk. Both cases are rare. Because schools are sometimes (but rarely) in a county other than the county of their district, we aggregate from the school-level to the county level. We obtain county FIPS codes from the CCD data when possible (when we can match schools between datasets) because the CRDC data does not contain county FIPS code for schools. In the rare cases that this is not possible, we rely on the ZIP code of the district (the CRDC data does not include addresses for schools and only contains ZIP, not county, code for districts). We obtain county code(s) for the ZIP code using a crosswalk (HUD PDR, 2017). In cases where the ZIP code contains only one county, we use that county. If it matches to multiple counties, we use an online geocoding tool from the Census[48] or information from NCES to obtain the county code (National Center for Education Statistics, 2020).

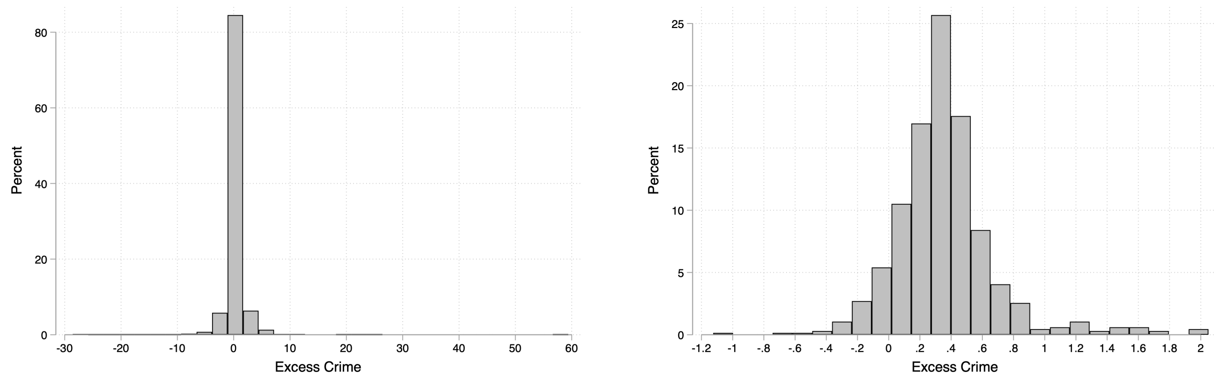
Constructing County-level Predictor Dataset

We first calculate excess crime at the county level. To do this, we aggregate reported crime to the county level, using the first-listed county FIPS code from the NIBRS data. We do not exclude counties that have police agencies that are associated with multiple counties. In addition to the sample restrictions from the overall excess crime calculation, we drop counties that do not have at least one observation for each of the 10–17, 19–24, 25–30, 31–35, and 36–40 age bins; or that do have at least one observation in July or between 8pm and 6am. We then calculate excess crime at the county level. After dropping additional counties that did not have enough information to compute a value, we merge on the covariates. We drop counties that are not matched (including counties with a '0' county code, which may include police agencies such as highway patrols) and that do not have nonmissing values for all covariates. We are left with 1,824 counties. We standardize all non-binary variables to have mean 0 and standard deviation 1. We present the distribution of excess crime in the Appendix Figure C.1. Panel A does not make restrictions; counties are centered around 0, with some big outliers. Panel B restricts to counties with at least 1,500 reported crimes among 10–40 year olds (after sample restrictions). These are centered around near the overall value of excess crime of 0.31. The extreme outliers are no longer present, presumably because the excess crime calculation is more imprecisely estimated with fewer observations. In our main specification, we weight by county population. We also present results in which we weight by the number of crimes. This gives counties with (presumably) more-precise estimates of excess crime more weight; these tend to be the larger counties (number of crimes has a correlation of 0.69 with county population). A disadvantage of this approach is that number of crimes is endogenous to the school year. We also present unweighted results and results weighted by grade 5–12

[48]https://geocoding.geo.census.gov/geocoder/geographies/onelineaddress/?form

enrollment.

Figure C.1: Distribution of County-level Excess Crime



(a) All Counties

(b) Counties with at Least 1500 Reported Crimes

Notes: This figure shows the distribution of county-level excess crime. Panel A shows all counties used in the analysis, while Panel B restricts to counties with at least 1,500 reported crimes in the data among 10–40 year olds, after excluding observations missing date, time, age, or occurring between midnight and 1am.