

Local Violence, Academic Performance, and School Accountability

By MARCUS CASEY, JEFFREY C. SCHIMAN, AND MACIEJ WACHALA *

* Casey: University of Illinois at Chicago, 601 S. Morgan (mc 144) Chicago, IL 60607 (email: mc Casey@uic.edu); Schiman: Georgia Southern University, P.O Box 8152 Statesboro, GA 30460 (email: jschiman@georgiasouthern.edu); Wachala: University of Illinois at Chicago, 601 S. Morgan (mc 144) Chicago, IL 60607 (email: mwacha2@uic.edu).

I. Introduction

Inspired by the “No Child Left Behind” legislation, school accountability systems seek to promote school quality improvements and close achievement gaps through increased scrutiny of schools and teachers. Standardized test scores and value-added measures largely determine “grades” assigned to institutions and teachers. Poor evaluations may have severe implications: teachers assigned bad grades may face pay decreases, discipline, and possibly, firing; in some cases, entire schools may be closed or replaced by a charter. Importantly, evaluation depends on within-school factors. However, a growing literature demonstrates that random shocks external to the classroom or school environment - for example, spikes in air pollution on testing days - may affect measured test performance (Ebenstein, Lavy, and Roth 2016). As such evaluation systems become increasingly attractive to districts, a more comprehensive understanding is needed

of how external stimuli may distort student test performance and, consequently, accountability ratings.

We study a salient shock: violent crime. Specifically, we match crime incidents to Chicago Public School (CPS) data to study how potential student exposure to local violent crime during testing periods may affect school-level performance and accountability scores. Chicago is ideal to study this issue with its large school system structured primarily around neighborhood schools. Moreover, its violent crime rate is relatively high, but unevenly distributed across neighborhoods. Hence, students attending demographically similar schools often face different potential violent crime exposure.

Our empirical strategy exploits variation in the timing and location of violent crime to study whether crimes occurring within testing periods near neighborhood schools are negatively related to student test scores and the probability of satisfying accountability standards. Our results suggest exposure to an additional violent event is associated with decreased test performance. The associations are quite local, dying out beyond 0.3 miles from the school. These

performance declines are consequential as the schools impacted by within-testing period violent crime are also less likely to meet accountability standards. Overall, our results suggest that violent crime exposure during testing periods may contribute to lower evaluations of schools.

Our paper adds to a literature demonstrating adverse effects of external shocks on students. Most closely related, Sharkey et al. (2014) find that Chicago school students exposed to local homicides exhibit lower vocabulary and reading scores. Similarly, Burdick–Will (2013, 2016) finds that violence on school campuses negatively affects student performance on cognitive tests. Studying a non-crime shock, Ebenstein et al. (2016) show that pollution levels affect student performance on high-stakes exams and that such effects may have lifelong implications. Our study complements this existing work by going beyond student-level outcomes and exploring the implications of such mechanisms for school evaluation systems.

II. Data and Empirical Model

Our data include test score and demographic information on all publicly-funded schools operating between 2004 and

2013 in Chicago.¹ We focus on neighborhood schools and construct school-level performance measures from the Illinois Standards Achievement Test (ISAT), information obtained from the Chicago Public Schools (CPS) Office of Accountability.² The ISAT is a standardized test of math and reading proficiency administered to 3rd through 8th graders each spring.³ Subject to some district-level guidelines, principals have discretion over test scheduling. Until the 2007–2008 academic year, the exam lasted six days; between 2009 to 2013, however, the testing period increased to ten days.

The crime data are drawn from the Chicago Police Department’s Citizen Law Enforcement Analysis and Reporting system. The crime database contains detailed information on the incidence of criminal activity between 2001 and 2014. They include the exact time and block where crimes occurred and the type: homicide, forcible rape, robbery, assault and battery, burglary, arson, and grand theft auto as well as simple property and quality-of-life crimes. We classify these crimes as violent or non-violent according to FBI definitions. Violent crime includes

¹ These schools include traditional public schools - neighborhood, magnet, and selective enrollment.

² We focus on neighborhood elementary schools because the student body is largely composed of children from the local community.

³ Elementary schools in Chicago typically include K-8.

homicide, forcible rape, robbery, and assault and battery. All other crimes are considered non-violent. We focus on violent crimes occurring within 0.3 miles of schools within their testing windows.⁴

In our empirical approach, we estimate variations of the following regression:

$$Outcome_{snt} = \sum_d \gamma_d [crime_{sdt}] + X_{st}\beta + \phi_{nt} + \epsilon_{snt}$$

Outcome denotes some accountability outcome at school *s*, in neighborhood *n* and year *t*. The vector X_{st} is a set of school tract-year controls, ϕ_{nt} denotes a neighborhood-by-time effect, and ϵ_{snt} represents unobservable determinants of school-level achievement; $crime_{sdt}$ denotes the incidence of crime within distance *d* of school *s* in year *t* during the testing week.

The parameters of interest, γ_d , measure how violent crime incidents during testing periods relate to school-level performance. Importantly, if violent crime shocks during the testing period are adversely associated with student performance, then failing to fully account for such shocks would wrongly

attribute poor performance or ratings to schools and teachers.

Note that identification of γ_d is potentially confounded by unobservable factors affecting student test performance that are correlated with the incidence of crime during testing week. Causal identification of γ_d requires the assumption that conditional on the inclusion of neighborhood/school and year fixed effects, violent crime incidents are as good as random.

III. Results

Table 1 presents results from equation (1) for several accountability-related outcomes: percentage of students exceeding or meeting their test growth goals (highest ISAT performance standard); percentage of students receiving an academic warning rating (lowest ISAT performance standard), and whether the school met Adequate Yearly Progress (AYP).⁵

⁵ During the sample period, students were tested using the ISAT. The state set various cutoffs measuring performance including meeting/exceeding; meeting; falling below; and academic warning (for examples of the criteria see <https://www.isbe.net/Pages/Performance-Definitions-ISAT-and-PSAE.aspx>, last accessed 1 October 2017). NCLB required states to set targets for attendance, test taking rates, and test growth. If the school met the targets, it would be deemed as making AYP. After several years of failing to make AYP, the school had to create and implement an action plan for improvement which would include changes like extending the school day or changing staff. Continued failure would ultimately result in a complete restructuring of school management. See <http://cps.edu/Programs/DistrictInitiatives/NCLB/Pages/NCLB.aspx> (last accessed 1 October 2017). After our sample period, the evaluation structure was adapted, but still based on criteria like attendance and test growth (see

⁴ In a descriptive statistics table available from the authors, we describe the average characteristics of schools in our sample and the distribution of violent incidents by distance.

For each outcome, we present models with different fixed effects including: (1) neighborhood and year fixed effects; (2) neighborhood-by-year fixed effects; or (3) school and year fixed effects. Overall, violent crimes occurring during testing periods are negatively associated with accountability metrics. The effect is quite local, dying out beyond 0.3 miles (about 2-3 city blocks). Panel A presents estimates of (1) where the outcome is the percentage of students that meet or exceed ISAT standards. Column (1) suggests that, conditional on neighborhood and year fixed effects, a violent crime during the testing period is associated with reductions in the percentage of students meeting or exceeding their goals. A similar pattern emerges in Column (2) where we allow for neighborhood-by-year fixed effects. When we condition on school-level fixed effects in Column (3), the estimates, though smaller, remain negative and statistically significant for crimes within 0.3 miles.

Panel B presents equation (1) for the percentage of students receiving an academic warning rating on the ISAT. In this case, violent crimes have statistically significant positive associations with the share of students receiving a warning. In particular, the estimates suggest that an additional crime

increases the share of students receiving a warning by 0.7 percentage points, a 5% increase relative to the mean. Again, the largest associations are for crimes occurring within 0.1 miles. Conditioning on school fixed effects reduces the estimates' magnitude, but they remain statistically significant. Taken together, these estimates suggest that higher local violent crime during testing raises the share of students receiving the lowest ISAT performance rating.

Panel C presents similar regressions where the dependent variable is the probability a school makes AYP. Consistent with panels A and B, each model suggests an additional violent crime is associated with reductions in the probability a school makes AYP by roughly 2 to 3 percentage points, around a 7% reduction relative to the mean. As with the prior estimates, the effect is strongest for crimes occurring within 0.1 miles of schools.

Our proposed mechanism is that violent crime exposure, which is highly disruptive, affects student performance. However, one might worry that the included fixed effects may not sufficiently capture all environmental factors that contribute to performance differences across schools. As supporting evidence, Table 2 presents estimates of the association of property crime on these outcomes. In general, property crime

is a nuisance and may reflect different dimensions of neighborhood quality. However, we would not expect property crime to change test performance and accountability metrics because property crime may be less salient than violent crime. Turning to Table 2, all but one of the estimates of property crime incidence on our outcomes are small and statistically insignificant. Altogether, these estimates support the notion that the impact of crime on student performance, and thus school-level evaluation, is concentrated on more salient violent crimes.

IV. Conclusion

This paper exploits variation in the location and timing of violent crime incidents with respect to neighborhood schools' annual testing periods. The evidence suggests that local violent shocks during testing periods are negatively associated with school-level test scores and school accountability standards. These results provide further evidence on the importance of incorporating localized shocks to student environments outside of school into the evaluation of schools and teachers.

REFERENCES

Burdick-Will, Julia. 2013. "School Violent Crime and Academic Achievement in Chicago."

Sociology of Education 86 (4).
doi:10.1177/0038040713494225.

———. 2016. "Neighborhood Violent Crime and Academic Growth in Chicago: Lasting Effects of Early Exposure." *Social Forces* 95 (1): 133–58. doi:10.1093/sf/sow041.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution." *American Economic Journal: Applied Economics*,8(4): 36-65.

Sharkey, Patrick, Amy Ellen Schwartz, Ingrid Gould Ellen, and Johanna Lacoë. 2014. "High Stakes in the Classroom, High Stakes on the Street: The Effects of Community Violence on Student's Standardized Test Performance." *Sociological Science* 1: 199–220.
doi:10.15195/v1.a14.

Table 1: Violent Crime Incidence During Testing Week and Measures of School Performance

Distance	Panel A. Exceeds or Meets			Panel B. Warning			Panel C. Meet AYP		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
0.1 miles	-0.010*** (0.003)	-0.011*** (0.003)	-0.002 (0.001)	0.008*** (0.002)	0.007*** (0.002)	0.005*** (0.001)	-0.028*** (0.007)	-0.024*** (0.008)	-0.019*** (0.006)
0.2 miles	-0.004*** (0.001)	-0.004** (0.002)	-0.002** (0.001)	0.004*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	-0.011*** (0.004)	-0.006 (0.005)	-0.009** (0.004)
0.3 miles	-0.001 (0.001)	-0.001 (0.001)	-0.001** (0.001)	0.002*** (0.001)	0.001 (0.001)	0.002*** (0.000)	-0.003 (0.003)	0.001 (0.003)	-0.004 (0.003)
Mean	0.415	0.415	0.415	0.142	0.142	0.142	0.316	0.316	0.316
NH FE	X			X			X		
Year FE	X		X	X		X	X		X
NH X Year FE		X			X			X	
School FE			X			X			X
N	3,662	3,662	3,662	3,662	3,662	3,662	3,662	3,662	3,662

Notes: Standard errors clustered by school are in parentheses. All regressions include controls for average per capita violent crime rates in the school-tract as well as violent crime counts before and after testing week. Regressions without school fixed effects also include school-tract demographic characteristics including unemployment by race, average household size, median household income, median age, and proportion below the poverty line. Mean refers to mean of outcome variable.

Table 2: Property Crime Incidence During Testing Week and Measures of School Performance

<u>Distance</u>	<u>Panel A. Exceeds or Meets</u>			<u>Panel B. Warning</u>			<u>Panel C. Meet AYP</u>		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
0.1 miles	-0.005 (0.003)	-0.005 (0.003)	-0.002 (0.001)	0.003 (0.002)	0.003 (0.002)	0.002 (0.001)	-0.006 (0.008)	-0.001 (0.009)	0.004 (0.007)
0.2 miles	-0.002 (0.001)	-0.002 (0.002)	-0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	-0.0017 (0.004)	-0.003 (0.004)	0.002 (0.004)
0.3 miles	-0.001 (0.001)	-0.001 (0.001)	-0.001* (0.001)	0.000 (0.000)	0.001 (0.001)	0.001 (0.000)	0.000 (0.002)	-0.000 (0.003)	0.002 (0.002)
Mean	0.415	0.415	0.415	0.142	0.142	0.142	0.316	0.316	0.316
NH FE	X			X			X		
Year FE	X		X	X		X	X		X
NH X Year FE		X			X			X	
School FE			X			X			X
N	3,662	3,662	3,662	3,662	3,662	3,662	3,662	3,662	3,662

Notes: Standard errors clustered by school are in parentheses. All regressions include controls for average per capita violent crime rates in the school-tract as well as violent crime counts before and after testing week. Regressions without school fixed effects also include school-tract demographic characteristics including unemployment by race, average household size, median household income, median age, and proportion below the poverty line. Mean refers to mean of outcome variable.