School Bus Diesel Emissions, Student Health, and Academic Performance: Evidence from Georgia

Wes Austin
Georgia State University

Garth Heutel
Georgia State University

Daniel Kreisman
Georgia State University

January 4, 2019

Motivation

A Major Contributor of Pollution

- 26 million American students ride over 500,000 school buses 4-6 billion miles per year. This represents about 2% of life between 5-17.

A Major Contributor of Pollution

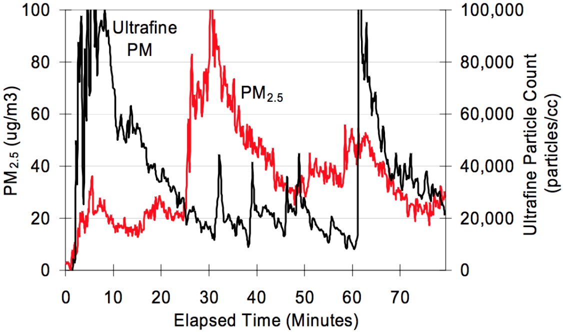


Figure 1: Conventional buses tested showed significant $PM_{2.5}$, ultrafine particle, black cabin and PAH self-pollution. (Ambient concentrations have been subtracted.)

Figure: Ultrafine and Fine PM in a Typical Bus Cabin

A Major Contributor of Pollution

Figure 3: Installation of diesel particulate filter on DeKalb bus prior to testing.

Figure: Making Dirty Buses Clean with Engine Retrofits

A Major Contributor of Pollution

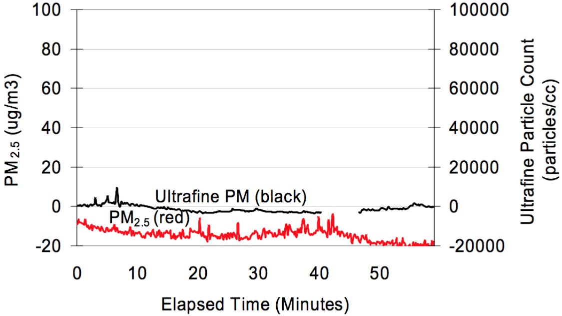


Figure 2: The DPF-ULSD-Spiracle combination eliminated $PM_{2.5}$, ultrafine particles, black carbon, and PAH self pollution from the bus cabin. Ambient concentrations have been subtracted resulting in slightly negative apparent net concentrations. Concentrations below zero should be taken as zero net $PM_{2.5}$ contribution to the bus.

Figure: Making Dirty Buses Clean with Engine Retrofits

A Major Contributor of Pollution

Figure: A Diesel Particulate Filter at the End of its Life

A Natural Experiment - School Bus Retrofits in GA

2,656 buses were retrofitted in 27 Georgia school districts from 2007-2015, affecting roughly 150,000 students.

- Average of 88 retrofits per grant, or 19% of the bus fleet.
- We can measure health (FitnessGram).
- We can test for achievement effects.
- We know the cost of the retrofit programs ($8,110 per bus).

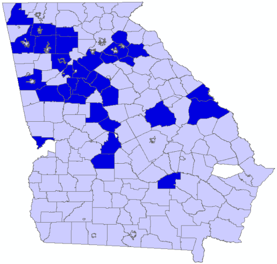


Figure: Retrofitting Districts (2007-2015)

Contribution

Bridges two literatures:

- We know air pollution is bad for human health and that it harms children more than adults (Stevens et al. 2010) (Currie et al. 2014) .
- We know that air pollution affects student achievement in the short- and long-term (Ebenstein et al. 2016) (Chen et al. 2017) as well as absences (Currie et al. 2009)
- Builds on the work of Beatty and Shimshack (2011) in Washington.

Research Methods

Data: Four Sources, District Level

1. **GA-DOE** (2007-2017)
   - Math/ELA test scores, student demographics.
2. **FitnessGram** (2012-2017)
   - PACER test of VO2 Max (aerobic capacity).
   - BMI, Curl-ups, Push-ups, and Sit and Reach
3. **GTA** (2010-2016)
   - Bus manifest - details of fleet and bus ridership.
4. **GADER** (2006-2016)
   - Retrofit type, number, month, year.

Summary Statistics

	(1) Non-Retrofitting		(2) Retrofitting		(3) T-Test of Means	
District Characteristics						
African American	0.367	(0.272)	0.363	(0.266)	0.004	(0.07)
Hispanic	0.082	(0.105)	0.109	(0.077)	-0.028	(-1.32)
White	0.554	(0.252)	0.504	(0.276)	0.051	(0.95)
FRL	0.668	(0.171)	0.616	(0.146)	0.052	(1.49)
SWD	0.123	(0.024)	0.121	(0.018)	0.002	(0.38)
ELL	0.025	(0.037)	0.045	(0.043)	-0.021*	(-2.58)
Students (thousands)	5.655	(9.765)	28.081	(37.502)	-22.426***	(-6.34)
District Bus Ridership	0.621	(0.174)	0.610	(0.087)	0.0113	(0.33)
Average Bus Age	14.126	(1.574)	14.268	(1.537)	-0.142	(-0.43)
Health and Education Outcomes						
Aerobic Cap. ($V0_2$)	41.160	(1.688)	41.201	(1.422)	-0.0412	(-0.12)
Body-Mass Index	21.069	(0.880)	20.633	(0.340)	0.436*	(2.54)
Math Z-Scores	-0.107	(0.263)	-0.060	(0.216)	-0.0473	(-0.88)
ELA Z-Scores	-0.107	(0.229)	-0.061	(0.194)	-0.0459	(-0.98)
Attendance rate	95.573	(0.630)	95.584	(0.488)	-0.0112	(-0.09)

First-Differences Research Model

We see retrofits, health and tests at the district level. Effects should be larger when a larger share of buses are retrofitted. For district i in year t:

$$\Delta y_{it} = \beta R_{it} + \Delta X_{it}\gamma' + \Delta Z_{it}\psi' + \tau_t + \Delta\epsilon_{it}. \tag{1}$$

- y_{it} is health (Aeroic Capacity, BMI) or schooling (ELA, Math, Attendance).
- R_{it} measures bus retrofits.
- X_{it} are district characteristics.
- Z_{it} are bus fleet characteristics.

Identification Assumption: There are no confounding factors correlated with the percent of a district bus fleet retrofitted and the change in health or education outcomes across all districts in their respective retrofit completion years.

Results

Health Outcomes

	Δ **Aerobic Capacity** (1)	Δ **BMI** (2)
% Retrofit	1.740**	-0.274
	(0.80)	(0.35)
Bus Chars.	✓	✓
Demog.	✓	✓
R2	0.189	0.053
N	856	856
n	180	180

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Column (1) suggests that if a district retrofitted its entire fleet, district-wide V02 max would increase by 1.74 ml/min/kg. With mean V02 Max of 41.2 ml/min/kg, column (1) suggests that if a district retrofitted its entire fleet, student aerobic capacity would improve by about 5%.

Aerobic Capacity by School Type

	Elementary		Middle		High School	
	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
% Retrofit	4.265**	4.241*	-1.772	0.336	1.874**	1.351
	(1.88)	(2.16)	(1.28)	(2.02)	(0.79)	(1.20)
Bus Chars.	✓	✓	✓	✓	✓	✓
Demog.	✓	✓	✓	✓	✓	✓
R2	0.148	0.275	0.095	0.305	0.032	0.086
N	776	776	767	767	710	710
n	173	173	170	170	168	168

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects and mean-centered ridership and trip duration variables included as controls. Columns (1) and (2) suggests that if a district retrofitted its entire fleet, elementary school V02 max would increase by about 10% of the mean.

Academic Outcomes

Next we look for test score differences:

- Now have data from 2007-2017.
- Same model with ELA and Math as dependent variables (z-scores).
- Check if attendance plays a role.

Academic Outcomes

	Δ **ELA Z-Score**	Δ **Math Z-Score**	Δ **Attendance**
% Retrofit	0.090***	0.047	0.183
	(0.03)	(0.03)	(0.25)
Bus Chars.	✓	✓	✓
Demog.	✓	✓	✓
R2	0.058	0.023	0.097
N	1,800	1,800	1,800
n	180	180	180

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects and mean-centered ridership and trip duration variables included as controls. For reference, the observed difference in elementary student test score gains from a teacher with five years of experience and a rookie teacher is 0.12 z-scores.

Academic Outcomes by School Type

	Elementary		Middle	
	ELA	Math	ELA	Math
	(1)	(2)	(3)	(4)
% Retrofitted	0.117***	0.057	0.060**	0.047
	(0.04)	(0.07)	(0.03)	(0.03)
R^2	0.043	0.02	0.042	0.038
N	1,800	1,800	1,800	1,800

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects and mean-centered ridership and trip duration variables included as controls.
Note: High-school students do not take ELA or Math end-of-grade tests in Georgia.

Summary of Findings

Health

- The mean aerobic capacity is 41 ml/min/kg, so retrofitting all buses in a district would lead to a 5% improvement in Aerobic Capacity.
- For ES students, who are more vulnerable to pollution and most likely to be tested, the effect is more than doubled.
- BMI placebo test suggests results are not driven by unrelated health changes.

Academic Outcomes:

- Modest gains in English (about 0.09 s.d. for retrofitting all buses).
- Positive but smaller and insignificant gains in Math.
- Does not seem to be driven by attendance, but mean is very high (0.95).

Policy Relevance

Costs

Costs: The average bus retrofit cost \$8,111 per bus in our sample.

- With 50 students per bus on average, this is \$160 per student.
- With a typical fleet size of 111 buses, retrofitting 10% of the typical district's fleet would cost \$90,000.

For a district transportation official, the alternative policy lever is to **buy new buses.**

- Regular diesel school buses cost \$130K but come with pollution control technologies.
- New hybrid and electric buses typically cost \$360K.
- Replacing 10% of the typical district's fleet would cost \$1.4M - \$4M.

Benefits

- Cardiovascular disease: A 1-MET increase in VO2max is associated with a 15% decrease in the lifetime risk of CVD.
  - Parent WTP for decreasing chance of CVD by 1 in 100 is \$139 over a child's discounted lifetime (Adamowicz et al. 2014).
  - Retrofitting 10% of its buses → 0.05 MET improvement → \$940,590 benefit.
- Test Scores: A one percentile increase in test scores is valued at \$1,041 over a working life after discounting (Chetty et al. 2011).
  - Z-score test results imply percentile increases of 0.33 and 0.16 in ELA and Math from retrofitting 10% of a bus fleet.
  - Taking the average (0.245) and multiplying by typical district size and the valuation of \$1,041 → \$2.3 million benefit.

Thank you!

Questions/comments?

Appendix

Robustness Concerns

Now we test some **alternative hypotheses**:

- Check if FitnessGram participation is changing (No).
- Check if bus ridership is changing (No).
- Drop milestones years (2015-17) → no change
- Exclude years with interpolation of bus variables (2007-2009 and 2016-2017) → no change
- Limit to elementary schools (effects are twice as large, BMI is negative).
- Test if method of assigning treatment years matters (possibly, although alternative treatment years increase magnitude and significance of math results)

Fitnessgram Participation and Ridership Robustness Check

	(1) AC attempts	(2) BMI attempts	(3) Ridership share
% Retrofit	-0.445	-0.576*	-0.023
	(0.32)	(0.30)	(0.04)
R2	0.149	0.030	0.012
N	870	870	1,780

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects, demographic and bus characteristics, and mean-centered ridership and trip duration variables included as controls. Each column is a separate regression.

Drop Interpolated Bus Years

	(1) ELA	(2) Math	(3) ATT	(4) AC	(5) BMI
% Retrofit	0.084***	0.056	0.283	1.687**	-0.274
	(0.03)	(0.04)	(0.29)	(0.83)	(0.36)
R2	0.079	0.029	0.176	0.179	0.065
N	1,260	1,260	1,260	692	698

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects, demographic and bus characteristics, and mean-centered ridership and trip duration variables included as controls. Each row-column cell is a separate regression.

Drop Milestones Years (2015-2017)

	(1) ELA	(2) Math
% Retrofit	0.089***	0.049
	(0.03)	(0.03)
R2	0.064	0.020
N	1,440	1,440

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects, demographic and bus characteristics, and mean-centered ridership and trip duration variables included as controls. Each row-column cell is a separate regression.

Timing of Retrofit Implementation

	(1) ELA	(2) Math	(3) Attend	(4) AC	(5) BMI
Regular Timing	0.090***	0.047	0.183	1.740**	-0.274
	(0.03)	(0.03)	(0.25)	(0.80)	(0.35)
Lagged Implementation	0.076***	0.073*	0.235	0.468	0.178
	(0.03)	(0.04)	(0.25)	(1.13)	(0.36)
January Implementation	0.100***	0.078**	0.299	1.649*	-0.114
	(0.02)	(0.03)	(0.26)	(0.85)	(0.29)
Drop Sept - April	0.132***	0.074**	0.062	2.879***	0.014
	(0.04)	(0.03)	(0.23)	(0.73)	(0.31)
1-year in Advance	-0.029	-0.040	-0.110	6.422***	0.197
	(0.03)	(0.04)	(0.23)	(2.26)	(0.62)
N	1,800	1,800	1,800	856	863
n	180	180	180	180	180

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects, demographic and bus characteristics, and mean-centered ridership and trip duration variables included as controls. Each row-column cell is a separate regression.

Results by Type of Retrofit

	(1) ELA	(2) Math	(3) Attend	(4) AC	(5) BMI
Diesel Particulate Filter	0.134**	0.063	0.459	1.411	-0.612
	(0.05)	(0.07)	(0.52)	(1.89)	(0.54)
Closed-Crankcase Filter	-0.022	-0.012	-0.635	-	-
	(0.04)	(0.05)	(0.45)	(.)	(.)
Diesel Oxidation Catalyst	0.051**	0.047	0.144	1.367	-0.139
	(0.02)	(0.03)	(0.19)	(0.85)	(0.46)
Flow-through Filter	-0.026	-0.177***	-0.149	-	-
	(0.06)	(0.05)	(1.43)	(.)	(.)
R2	0.058	0.023	0.096	0.186	0.049
N	1,800	1,800	1,800	856	863
n	180	180	180	180	180

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects, demographic and bus characteristics, and mean-centered ridership and trip duration variables included as controls. Each row-column cell is a separate regression.

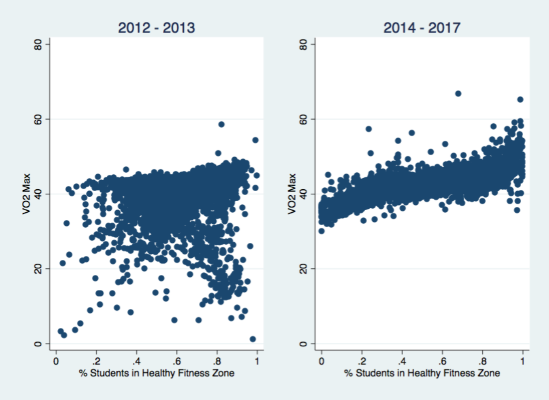
Academic Results by Grade

	(1) ELA	(2) Math
3rd Grade	0.092	0.036
	(0.07)	(0.11)
4th Grade	0.213**	0.204*
	(0.10)	(0.11)
5th Grade	0.158***	0.047
	(0.05)	(0.08)
6th Grade	0.053	-0.001
	(0.05)	(0.06)
7th Grade	0.059	0.041
	(0.05)	(0.04)
8th Grade	0.063	0.102
	(0.04)	(0.08)
N	1,800	1,800
n	180	180

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the district level. Year fixed effects, demographic and bus characteristics, and mean-centered ridership and trip duration variables included as controls. Each row-column cell is a separate regression.

Imprecision in Aerobic Capacity

Figure: Imprecision in Aerobic Capacity Measurements



Average aerobic capacity in 2012 is left-skewed, with many implausibly low values for $V0_2Max$. In later years, no school-average $V0_2Max$ is below 30 for male assessments and 26 for female assessments.

Imprecision in Aerobic Capacity

	(1)	(2)	(3)	(4)
	15	20	25	30
% Retrofit	2.327	1.324	1.507	1.760**
	(2.05)	(1.33)	(0.95)	(0.73)
R2	0.247	0.238	0.223	0.218
N	860	860	860	857

	(5)	(6)	(7)	(8)	(9)
	35	30 & 26	Jumps	None	2012
% Retrofit	1.637**	1.740**	3.445*	1.368	7.133***
	(0.73)	(0.80)	(1.99)	(2.26)	(1.14)
R2	0.149	0.189	0.267	0.246	0.300
N	849	856	638	860	681

Columns (1)-(5) are aerobic capacity results when dropping school observations below the given value. In column (6) we drop school observations below 30 for males and below 26 for females. In column (7) we replace as missing any school with average values that increase or decrease by more than 6 $V0_2Max$ units from 2011-12 to 2012-13. In column (9) we drop the entire year of 2012. We prefer model (6), the cutoff at 30 for males and 26 for females.