Syracuse University

SURFACE at Syracuse University

Center for Policy Research

Maxwell School of Citizenship and Public Affairs

12-2021

Using Pupil Transportation Data to Explore Educational Inequities and Outcomes: A Case Study from New York City

Sarah Cordes
Temple University, sarah.cordes@temple.edu

samantha Trajkovski
Syracuse University, setrajko@syr.edu

Christopher Rick
Syracuse University, carick@syr.edu

Meryle Weinstein
New York University, meryle.weinstein@nyu.edu

Amy Ellen Schwartz
Syracuse University, amyschwartz@syr.edu

Follow this and additional works at: https://surface.syr.edu/cpr

 Part of the Economic Policy Commons, and the Economics Commons

Recommended Citation

Cordes, Sarah; Trajkovski, samantha; Rick, Christopher; Weinstein, Meryle; and Schwartz, Amy Ellen, "Using Pupil Transportation Data to Explore Educational Inequities and Outcomes: A Case Study from New York City" (2021). *Center for Policy Research*. 376.
https://surface.syr.edu/cpr/376

This Working Paper is brought to you for free and open access by the Maxwell School of Citizenship and Public Affairs at SURFACE at Syracuse University. It has been accepted for inclusion in Center for Policy Research by an authorized administrator of SURFACE at Syracuse University. For more information, please contact surface@syr.edu.

CENTER FOR POLICY RESEARCH
THE MAXWELL SCHOOL
WORKING PAPER SERIES

Using Pupil Transportation Data to Explore Educational Inequities and Outcomes: A Case Study from New York City

Sarah Cordes, Samantha Trajkovski, Christopher Rick, Meryle Weinstein, and Amy Ellen Schwartz

Paper No. 243
December 2021

ISSN: 1525-3066
426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
T 315.443.3114 **E** ctrpol@syr.edu
https://www.maxwell.syr.edu/CPR_Working_Papers.aspx

Syracuse University
Maxwell School of
Citizenship & Public Affairs
Center for Policy Research

CENTER FOR POLICY RESEARCH - Fall 2021

Leonard M. Lopoo, Director
Professor of Public Administration and International Affairs (PAIA)

Associate Directors

Margaret Austin
Associate Director, Budget and Administration

John Yinger
Trustee Professor of Economics (ECON) and Public Administration and International Affairs (PAIA)
Associate Director, Center for Policy Research

SENIOR RESEARCH ASSOCIATES

Badi Baltagi, ECON
Robert Bifulco, PAIA
Carmen Carrión-Flores, ECON
Sean Drake, SOC
Alfonso Flores-Lagunes, ECON
Sarah Hamersma, PAIA
Madonna Harrington Meyer, SOC
Colleen Heflin, PAIA
William Horrace, ECON
Yilin Hou, PAIA
Hugo Jales, ECON
Jeffrey Kubik, ECON
Yoonseok Lee, ECON
Amy Lutz, SOC
Yingyi Ma, SOC
Jerry Miner, ECON
Shannon Monnat, SOC
Jan Ondrich, ECON
David Popp, PAIA
Stuart Rosenthal, ECON
Michah Rothbart, PAIA
Alexander Rothenberg, ECON
Rebecca Schewe, SOC
Amy Ellen Schwartz, PAIA/ECON
Ying Shi, PAIA
Saba Siddiki, PAIA
Perry Singleton, ECON
Yulong Wang, ECON
Peter Wilcoxen, PAIA
Maria Zhu, ECON

GRADUATE ASSOCIATES

Rhea Acuña, PAIA
Graham Ambrose, PAIA
Mariah Brennan, SOC. SCI.
Brandon Charles, PAIA
Ziqiao Chen, PAIA
Yoon Jung Choi, PAIA
Stephanie Coffey, ECON
Adam Cucchiara, PAIA
William Clay Fannin, PAIA
Giuseppe Germinario, ECON
Myriam Gregoire-Zawilski, PAIA
Joshua Grove, SOC
Hyojeong Kim, PAIA
Mattie Mackenzie-Liu, PAIA
Maeve Maloney, ECON
Austin McNeill Brown, SOC. SCI.
Qasim Mehdi, PAIA
Nicholas Oesterling, PAIA
Claire Pendergrast, SOC
Lauryn Quick, PAIA
Michael Quinn, ECON
Radine Rafols, ECON
Christopher Rick, PAIA
Sam Saltmarsh, ECON
Spencer Shanholtz, PAIA
Sarah Souders, PAIA
Yue Sun, SOC
Joaquin Urrego, ECON
Yao Wang, ECON
Zhanhan Yu, ECON
Xiaoyan Zhang, Human Dev.
Bo Zheng, PAIA
Dongmei Zuo, SOC. SCI.

STAFF

Willy Chen, Research Associate
Michael Dunaway, Postdoctoral Scholar
Katrina Fiacchi, Administrative Specialist
Michelle Kincaid, Senior Associate, Maxwell X Lab
Emily Minnoe, Administrative Assistant
Hannah Patnaik, Managing Director, Maxwell X Lab
Candi Patterson, Computer Consultant
Xiaohan Sun, Postdoctoral Scholar
Samantha Trajkovski, Postdoctoral Scholar
Laura Walsh, Administrative Assistant

Abstract

This article explores how researchers can use pupil transportation data to explore key questions about the role of transportation in educational access and equity, such as how students get to school and the effect of transportation on student outcomes. We first describe different sources of transportation data that are available to researchers, provide a brief review of relevant literature, and discuss potential sources of measurement error in pupil transportation data. Next, we use administrative data from New York City to illustrate how pupil transportation data can be used to understand transportation eligibility and assignment as well as to describe the characteristics of students' commutes to school. For example, we find that not all students assigned for free transportation take it up. Specifically, although 47 percent of K-12 students in 2017 were eligible for pupil transportation based on distance with another 9 percent of students receiving exceptions, only 45 percent of students were assigned to a full-fare MetroCard, general education bus, or special education bus. Further, we find the average commute to school for walkers and bus riders is quite similar—around 30 minutes—although there is wide variation as some students experience very short or very long commutes. We end with a discussion of the importance of the institutional context when conducting research using pupil transportation data and best practices when using administrative data.

JEL No.: I20, I24, I29

Keywords: Education, Pupil Transportation, School Bus, Commuting

Authors: Sarah Cordes, Temple University, sarah.cordes@temple.edu; Samantha Trajkovski, Center for Policy Research, Syracuse University, setrajko@syr.edu; Christopher Rick, Center for Policy Research, Department of Public Administration and International Affairs, Syracuse University, carick@syr.edu; Meryle Weinstein, New York University, meryle.weinstein@nyu.edu; Amy Ellen Schwartz, Center for Policy Research, Department of Economics, Syracuse University, amyschwartz@syr.edu

Acknowledgements

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through GrantR305A170270 awarded to Syracuse University (Schwartz). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

We thank the NYCDOE Office of Pupil Transportation for data access and their continued support for this project and Joanna Bailey and Willy Chen for their research assistance on the project.

Contents

1. Pupil Transportation Data Basics
- 1.1. What can we learn from pupil transportation data?
- 1.2. Sources of Pupil Transportation Data
  - 1.2.1. National Transportation Data
  - 1.2.2. District-Level Administrative Pupil Transportation Data
  - 1.2.3. Example: New York City Pupil Transportation Data
    - 1.2.3.1. Variables Provided by OPT
    - 1.2.3.2. Variables Calculated from OPT Data

2. Transportation Research
- 2.1. ECLS-K Studies
- 2.2. Single District Research Studies Using Pupil Transportation Data
- 2.3. Studies Using Estimated Eligibility and Transit Routes

3. Potential Sources of Measurement Error
- 3.1. Coverage
- 3.2. Transportation Assignment and Use
- 3.3. Commute Times

4. **NYC Case Study**
- 4.1. Overview and Institutional Context
  - 4.1.1. Eligibility
  - 4.1.2. Assignment
  - 4.1.3. Exceptions
- 4.2. Who is eligible for and assigned to pupil transportation?
- 4.3. What are the characteristics of student commutes?

5. Special Issues with Transportation Data
- 5.1. Institutional Context
  - 5.1.1. Transportation Routing and Management
  - 5.1.2. Routing Restrictions
  - 5.1.3. Eligibility and Assignment
  - 5.1.4. Transportation Exceptions
  - 5.1.5. Other Modes of Transportation
  - 5.1.6. Charter and Choice Schools
- 5.2. Best Practices for Using Administrative Data for Research
  - 5.2.1. Data Privacy, Security, and Management
  - 5.2.2. Data Linking

Appendix

1. Pupil Transportation Data Basics

Pupil transportation data, at its core, describes *how* students get to school. This can include student-level information about bus route assignments or eligibility for bus service, as well as school and district-level information about spending. While pupil transportation is fundamental for operating schools, transportation has been largely ignored in the education policy arena. However, a growing literature is beginning to examine the role of transportation in improving student success, partly due to increasing availability of pupil transportation data. This data guide provides an overview of pupil transportation data and the role that transportation plays in students' lives, using longitudinal student-level pupil transportation data from New York City.

1.1 What can we learn from pupil transportation data?

Despite the ubiquity of the yellow school bus in depictions of education, the research on the role of pupil transportation remains limited. This provides an excellent opportunity for scholars, as numerous important and interesting research questions about pupil transportation have yet to be answered. Given how little we know about pupil transportation, there is a need for good descriptive work. This work could focus on how students get to school (mode of transportation), how long it takes them to get there, the variation in commute times, and the composition of the school bus (race/ethnicity, age, etc.). There are also important causal questions to be explored, for example, how commute time and mode of transportation affects student outcomes, the role of transportation in school choice, and how different forms of transportation affect student engagement and participation in after school activities.

1.2 Sources of Pupil Transportation Data

1.2.1 National Transportation Data

There are no detailed, national data focused on pupil transportation. National student-level transportation data are typically collected as part of larger surveys. For example, the Early Childhood Longitudinal Survey - Kindergarten Class of 1999 (ECLS-K:99), includes a question on the parent survey that asks how students get to school. While this is helpful to determine general patterns of transportation use nationally, it does not allow researchers to explore more nuanced questions regarding commute times or causal questions. Furthermore, this data is only available for a single year, and may not reflect current commuting patterns.

The National Household Travel Survey (NHTS) collects individual-level data on travel and transportation behavior and includes information on trips made by all modes of travel (e.g., private vehicle, public transportation) and for all purposes (e.g., travel to work, school, etc.). NHTS data has been conducted periodically since 1969, with the most recent waves collected in 2009 and 2017. Each wave of the survey includes travel decisions for over 30,000 school children, which allows researchers to examine changes in student commute patterns over time. For example, in 2017, 36.5 percent of students reported taking the school bus, while 50.2 percent were driven in a private vehicle. An additional 1.9 percent of students commuted to school using public buses or subways. This is a sizeable increase in the share of students commuting to school in private vehicle since the first survey in 1969, when only about 15 percent of students commuted by private vehicle. While these data are useful for describing trends in pupil transportation over time, because the survey is only offered periodically, does not track the same students over time, and does not contain information on other student characteristics and outcomes,

these data are limited in their capacity to explore inequities in transportation or the role of transportation in shaping student outcomes.

District-level spending data are available from the Local Education Agency School District Finance Survey Data collected by the National Center for Education Statistics. These data allow researchers to examine changes in transportation spending over time, examine disparities in transportation spending across districts, and, in combination with data from the Local Education Agency University Survey data, examine correlations between district demographics and transportation spending.

1.2.2 District-Level Administrative Pupil Transportation Data

Since pupil transportation is typically managed at the district level, another source for detailed pupil transportation data that can be linked with other student-level information and outcomes is the district office responsible for pupil transportation. While specific data availability will differ from district-to-district, there are a set of key variables that researchers interested in pupil transportation for education should request and which may be available from the district since they are needed by district personnel for routing students to school.

These key variables are:

- Student IDs – this is necessary to link transportation data to other records
- Transportation eligibility or data necessary to determine eligibility
    - Type of eligibility: school bus, subsidized public transportation, etc.
    - Reason for eligibility: distance from school, required by IEP, living in homeless shelter, etc.
- Assignment – what transportation does student *receive*

    - Form of transportation: school bus or public transportation
    - General versus special education bus
- Characteristics of service
    - AM and PM bus stop locations
    - AM and PM pick-up/drop-off times
- Other
    - Route IDs -allows the researcher to identify students who ride the same bus
    - Exceptions: reasons otherwise ineligible student receives transportation, such as living in a homeless shelter.
    - School start times

These data should be able to be linked with other data containing relevant student characteristics and outcomes such as home address coordinates, student demographics, performance on standardized tests, etc.

1.2.3 Example: New York City Pupil Transportation Data

Pupil Transportation data in New York City (NYC) are collected by the Office of Pupil Transportation (OPT) at the NYC Department of Education (NYCDOE). They are shared with researchers through a Data Request to the Research and Policy Support Group (RPSG). Once this request is approved, a non-disclosure agreement is signed between the NYCDOE and the requesting institution.

The following table lists the variables we use in our work on pupil transportation in NYC. Table 1 describes variables provided by directly OPT and which may be available in other districts and Table 2 describes variables that we calculate.

1.2.1.1 Variables Provided by OPT

The data provided by OPT includes the variables that are used to determine transportation eligibility and those generated through the assignment and routing process.

Table 1: Variables in NYC Pupil Transportation Data

Variable Name	Definition
Provided By OPT	
Distance zone	Shortest street-network distance between home and school as calculated by NYC OPT
	A - Less than 0.5 miles
	B - Between 0.5 and 1.0 miles
	C - Between 1 and 1.5 miles
	D - Greater than 1.5 miles
Transportation Assignment	Transportation option assigned to the student
	Half-fare MetroCard
	Full-fare MetroCard
	General education school bus
	Special education school bus
	No assignment
Exception Code	Code indicating reason for transportation exception. Missing values indicate no exception.
AM Route	Unique bus route ID for morning transportation
PM Route	Unique bus route ID for afternoon transportation
Pick Up Time	Expected pick-up time at AM bus stop
Drop Off time	Expected drop-off time at PM bus stop
Start Time	School start time
End Time	School end time

1.2.1.2 Variables Calculated from OPT Data

Based on the data from OPT, we calculate a set of variables capturing transportation eligibility, exceptions, and characteristics of the commute. Unless otherwise noted, all commute distances are calculated using the *Open Source Routing Machine* (OSRM). OSRM uses geographic data on latitude and longitude to determine travel time and street network distance between two coordinate pairs using

a user-imported map of any city from *OpenStreetMaps*. To estimate commute times for bus riders, we use HERE Location Services. HERE is an application interface that provides real-time routing between pairs of coordinates. The application finds the fastest travel route between pairs of coordinates and returns the travel time with traffic, the travel time without traffic, and the distance. In our work, we use estimates of driving time with traffic at 7:30AM on a weekday morning in May 2019 when there were no major incidents or crashes that would overstate travel times. All distances and travel times originating at a student's home are calculated using home address coordinates, which are provided by the RPSG under a different data request. Lastly, because whether to offer the bus is a school-level decision, we use student-level data to create a school-level measure of bus availability.

Table 2. Variables Calculated from OPT Data

Calculated	
Transportation Eligibility	Transportation option that the student is eligible for based on distance zone and grade level
	Not eligible
	Half-fare
	Full-fare or school bus
	Full-fare only
Transportation Exception	= 1 if student received any transportation exception
Walk distance from home to school	Minimum street network distance between home and school, in miles
Drive distance from home to school	Minimum driving distance between home and school based on street network, in miles
Walk distance from home to bus stop	Minimum street network distance between home and bus stop
Walk time from home to school	Estimated minimum walk time between home and school, in minutes
Drive time from home to school	Estimated drive time between home and school based on street network distance and traffic, in minutes
Walk time from home to bus stop	Estimated minimum walk time between home and AM bus stop, in minutes
Time from morning pick up to last stop	Minutes between AM pick up and the last stop on the route

Calculated	
Time from last stop to drop off	Estimated drive time between last stop on AM route and arrival at school based on street network distance and traffic, in minutes
Time from pick up to drop off	Minutes between AM pick up time and estimated school drop off time
Time from drop off to start time	Minutes between estimated school drop off and start time
Time from morning pick up to school start time	Minutes between AM pick up time and school start time
Time from end of school to afternoon drop off	Minutes between school end time and PM drop off time
Time from morning pick up to afternoon drop off	Hours between AM pick up time and PM drop-off time
Student attends school which offers a bus	=1 if 5 or more students at the school is assigned a bus and at least one of these does not have a transportation exception

2. Transportation Research

2.1 Studies using ECLS-K

Two recent papers use national data from the ECSL-K to explore the relationship between use of the school bus and absenteeism. Gottfried (2017) finds that kindergartners taking the bus are absent about 0.4 fewer days per year, a statistically significant difference that may have small, but potentially meaningful effects on academic outcomes. More recently, Gottfried et al (2021) find kindergarten school bus riders in rural communities miss fewer school days than bus riders in suburban, urban, or town communities and are less likely to be chronically absent.

2.2 Single District Research Studies Using Pupil Transportation Data

A small number of students leverage pupil transportation data, like that described above, to explore how transportation eligibility and assignment is related to attendance and school choice.

Cordes et al. (2019) use New York City student-level administrative data on bus ridership and absenteeism to examine the link between bus riding and attendance rates. They find that bus riders are absent about one fewer day per year than peers that do not ride the bus, and that much of the absenteeism gap for bus riders is driven by differences between, rather than within, schools. They also find that the absenteeism gap for Black (1.8 fewer days absent) and Hispanic (1.4 days) bus riders is three to four times larger than for Asian (0.4 days) students and 20 to 25 times larger than for White (0.1 days) students.

Two papers examine the link between pupil transportation and school choice. Cordes and Schwartz (2018) examine the relationship between transportation and school choice for elementary school students in New York City. They find bus riders are more likely to attend a choice school (rather than their zoned school). And, among students who attend choice schools, those taking the bus attend significantly better schools than those who commute on foot or rely upon another form of transportation.

More recently, Trajkovski et al (2021) explores the role of distance and transportation, namely the school bus, on school choice decisions for NYC kindergarten students in 2017. Results suggest that while students are less likely to choose a school farther from home, school bus eligibility increases the likelihood of attending a school by 1.4 to 4 percentage points. Further, they find significant heterogeneity by school type, with the largest effect for charter schools; bus eligibility reduces the negative effect of distance by 68-99 percent.

2.3 Studies Using Estimated Eligibility and Transit Routes

Recently, a number of studies have also explored issues of transportation focusing on how transportation eligibility, travel time, and public transit route characteristics may affect students. Rather than use pupil

transportation data, these studies *estimate* transportation eligibility or likely routes to school using data on students' location of residence and school attended and district policies regarding eligibility.

Edwards (2021) examines the effect of pupil transportation eligibility on attendance and achievement in for K-8 students in Michigan using a regression discontinuity design. She finds small positive (negative) effects of transportation eligibility on attendance (chronic absenteeism). Further, the effects are larger for economically disadvantaged students, where results suggest eligibility increases attendance by 0.5-1 percentage points and reduces chronic absenteeism by 2-4 percentage points.

Pogodzinski et al (2021) also explore transportation eligibility and attendance in Detroit, where they find a small negative association between transportation eligibility and attendance. They argue that these findings, although counterintuitive, may indicate that transportation eligibility does not always translate into regular transportation-use and that eligibility alone may not be enough to overcome negative relationships between student poverty and attendance.

Blagg, et al. (2018) use student-level data from Washington, DC to compare test scores and attendance rates for students with long and short commutes by car or public transit both between and within schools. They find that students with commutes estimated at the 75th percentile (14 minutes) are absent about one more school day per year than classmates with commutes estimated at the 25th percentile (3 minutes). Despite increased rates of absenteeism, they find no difference in test scores for students with longer commutes. This could indicate that students with longer commutes have unobserved characteristics, like motivation, that lead them to commute further to school and enable them to score similarly on standardized tests, despite higher absenteeism rates.

Finally, Stein and Grigg (2019) and Stein et al (2021) examine the relationship between features of the commute such as number of vehicles used, travel distance, and number of transfers and student absenteeism and school transfers in Baltimore. Specifically, Stein and Grigg (2019) find that increased travel time between 8th and 9th grade is associated with higher rates of absenteeism in 9th grade, while Stein et al (2021) find that increased travel time is associated with a higher likelihood of high school transfer. Further, among students who transfer schools, they tend to enroll in schools that are closer to home and are less likely to have been ranked highly in their initial high school choice application.

2.4 Conclusion

Taken together, these studies illustrate the potential for pupil transportation data to address important questions regarding student outcomes, school access, and school choice. They also demonstrate that there are many important questions still unanswered.

3. Potential Sources of Measurement Error in Pupil Transportation Data

3.1 Coverage

There main source of coverage error in pupil transportation data is that not all students may be included. This will primarily be a concern if pupil transportation data are "point-in-time," as students who are not enrolled at the time the data is pulled will be excluded. This may be exceptionally problematic for mobile students who move in and out of the school district. In NYC, this problem is relatively small, however, as we are able to successfully link 95.81% of our student-level observations with transportation data.

First, not all students are included in the data. For example, students who are not assigned transportation are excluded from the data. We are, therefore, unable to draw conclusions on students who come to school by car, walk, or bicycle and limits our ability to compare the effectiveness of different forms of

transportation. It may also lead underestimates of the effects of transportation to school. In NYC, approximately 39 percent of students are not assigned a transportation mode.

3.2 Transportation Assignment and Use

There is also likely to be measurement error in students' transportation assignment, as data provided by district offices will reflect students' assignments at the time that data are pulled. Since students may move homes, switch schools, or opt out of transportation throughout the school year, the data provided may not accurately capture students' assignments at the time that outcomes are measured or throughout the year. This may be especially problematic for students who move in and out of homeless shelters or are moved to a different shelter throughout the year. For example, in NYC, what we observe with these point-in-time data is likely an underestimate of the total number of homeless students who receive pupil transportation throughout the entire school year. This limits our ability to explore the heterogeneous impacts of transportation by time of year or weather conditions.

A further limitation is that researchers may not be able to identify transportation mode for those students who are assigned no transportation. For example, in NYC, we are unable to determine whether these students walk, bike, are driven, or use some other mode of transportation to get to school.

Finally, while the district may provide information on transportation *assignment*, they may not report *use*. While it is true that students may be unassigned from the bus after frequent disuse, for this to occur, it must be reported to the school. For students who are assigned to bus stops with multiple other riders, this is unlikely to occur because the bus driver may not notice if any particular student is missing. Therefore, even with data on transportation assignment, researchers may incorrectly attribute bus use to students who opt for other forms of transportation and there may be unobserved selection into or out of bus service, if for example, students decide to forgo bus service due to a long bus ride

3.3 Commute Times

While districts may have data on bus pick-up times, as these are generated for routing purposes, they may not have GPS data that reflect actual pick-up and drop-off. Therefore, researchers may need to estimate commute distance and time using mapping and routing software. This introduces measurement error that may lead researchers to underestimate the effects of commutes on student outcomes and may also mask considerable heterogeneity in commute times due to differences in traffic. However, NYC is making efforts to equip school buses with GPS to track students and bus location, and we may be able to receive these nuanced data in future years.

Lastly, for students using public transportation, districts are unlikely to have or be able to provide information about commutes or travel times. While routes can be estimated using data on students' residential addresses and mapping software such as Google Maps, ultimately researchers do not know what routes these students take to school or how long their commutes are. Further, it is challenging to account for systemic issues, such as train delays or weather-related events that could delay trains and public buses. This again hampers the ability of researchers draw a causal link between commute times and student outcomes. In the future, we hope to address some of these issues in NYC by requesting MetroCard barcodes to link with swipe data from the MTA, which would allow us to determine where students enter the system in the mornings and afternoon and better approximate their travel route.

4. NYC Case Study

In this section, we present more details on the transportation data we use in NYC to provide an example of how pupil transportation data can be used to understand the distribution of students eligible, and assigned to, pupil transportation and the characteristics of student commutes. Details about the data cleaning process are included in the appendix.

4.1 Overview and Institutional Context

As noted earlier, data on student transportation for NYC public school students is provided by the NYCDOE OPT. These student-level data includes information on whether students are *eligible* to receive free/subsidized transportation to school, their *assigned* transportation option (MetroCard or yellow school bus), and a variety of characteristics about individual bus routes. The data is available for all general education students in grades K-12 and students in ungraded special education. To explore questions around transportation equity and impacts, we link these to other data including student demographic and program information, performance, attendance, and home address coordinates, which are also obtained from RPSG.

4.1.1 Eligibility

Eligibility for general education students is determined by a student's grade and how far he or she lives from school. Table 3 presents the distance-based eligibility rules as of September 2019. Students in grades K-2 must live at least a half mile from school to be eligible for school bus service or a full-fare MetroCard; this distance increases to 1 mile for students in grades 3-6. After grade 6, students are no longer eligible for the bus, except in Staten Island, where general education bus service is offered through 8th grade because of lack of access to subways.

Table 3: Transportation Assistance Eligibility Rules

	Distance between Home and School			
Grade	A) <.5 miles	B) .5-1 mile	C) 1-1.5 miles	D) ≥1.5 miles
K-2	Half-Fare MC	Full-Fare MC or School Bus	Full-Fare MC or School Bus	Full-Fare MC or School Bus
3-6	Not Eligible	Half-Fare MC	Full-Fare MC or School Bus	Full-Fare MC or School Bus
7-12	Not Eligible	Half-Fare MC	Half-Fare MC	Full-Fare MC

4.1.2 Assignment

Once students are deemed eligible for transportation assistance, they may receive one of the following:

Half-Fare MetroCards

Students in grades K-2 (3-6) who live less than a half mile (between a half and 1 mile) from school are eligible for a half-fare MetroCard. These are public transit passes that allow students to ride city buses for half the regular fare, with the other half paid by students' families. Students who *eligible* for half-fare MetroCards must request them, but many not do so because the benefits are relatively small. As of AY 21 half-fare MetroCards have been phased out.

Full-Fare MetroCards

Students in grades K-2 (3-6) who live *more* than a half mile (1 mile) are eligible for a full-fare MetroCard, which can be used up to three times a day on either the subway or bus, Monday through Friday when school is in session. Students in grades 7-12 must live more than 1.5 miles from school in order to be eligible for a full-fare MetroCard. Students who are *eligible* for MetroCards do not need to request them but may decline the offer if they decide they would rather rely on some other form of private transportation, such as private cars or taxis (in NYC this includes ride shares such as Uber and Lyft, among others).

School Bus Transportation

The yellow school bus picks students up from a designated stop and drops them off at school in the morning (stop-to-school service), and the reverse in the afternoon. Students who ride the bus are responsible for getting to and from their bus stops.

To receive school bus service, students must meet three conditions. First, they must meet distance eligibility criteria, which are the same as the criteria to receive a full-fare MetroCard, or be given an exception. Second, eligible students must attend a school that offers the bus. Since each individual principal decides whether or not to offer bus service, some eligible students may not receive bus service because it is not offered by their school. Third, OPT must be able to accommodate the student on a route that meets the following conditions:

- the route cannot be longer than five miles from first stop to the school
- the route must have at least 11 students at the time of creation
- the route cannot cross community school district boundaries (traditional public schools) or borough boundaries (charter schools).

Therefore, even if a student attends a school that offers bus service, that student may not receive bus service due to an inability to accommodate them on a route that meets these restrictions.[1]

Based on these restrictions, administrators at schools that offer bus service then request either the school bus or full-fare MetroCards for each eligible student. Administrators also then assign bus stops to those students for whom they are requesting bus service. These assignments are then verified by OPT to ensure that they meet routing restrictions. As with MetroCards, students who are offered school bus service may choose to eschew it in favor of other forms of transportation.

Special Education

Some special education students are eligible for door-to-door service because it is mandated by their individualized education plans (IEP). As suggested by the name, students receiving door-to-door service are picked up at their home and dropped off at school. Depending on the nature of the transportation mandate in their IEP, some special education students may instead receive service on the general education bus. Transportation eligibility and assignment for special education students without an IEP mandate is determined by the same criteria governing general education students.

4.1.3 Exceptions

There are numerous exceptions that extend transportation access to students who might otherwise not be eligible. These exceptions are offered for a variety of reasons, including injury, hazardous walking routes, or because their zoned school is designated as failing under NCLB and they are attending another school under the transfer option. One of the most common exceptions is for students living in homeless

[1] Elementary schools, charter schools, and schools with gifted and talented programs are more likely to provide buses, as are schools in lower density areas. The probability a school offers a bus is increasing in the share of the student body that is White or Asian, eligible for pupil transportation or live outside the attendance zone. For more information on the characteristics of schools that offer the bus in NYC, see Cordes et al. (2020).

shelters, who are automatically eligible for either the school bus or full-fare MetroCards under the McKinney-Vento Act. For many years, these students were only assigned a bus if they could be accommodated on an existing route. As of 2016, however, a policy change by the mayor mandated that all students in homeless shelters be provided a school bus unless they opt-out to receive a full-fare MetroCard.

4.2 Who is eligible for and assigned to pupil transportation?

In 2017, 20 percent of students were eligible for the school bus or full-fare MetroCard, while an additional 27 percent of students were eligible for a full-fare MetroCard (Table 4). In total, then, almost half (47 percent) of K-12 students were eligible for some form of free transportation based on how far they live from school. Another 9 percent of students received exceptions. Not all students who were eligible for free transportation took it up, however, as only 45 percent of students are assigned to full-fare MetroCards (34 percent), a general education bus (7 percent), or a special education bus (4 percent). There is also particularly low take-up of half-fare MetroCards—although 33 percent of students are eligible based on distance, only 12 percent of students are assigned to this form of transportation.

Table 4: Summary Statistics for NYC Pupil Transportation Data, 2017

Variable Name	Observations	Mean	Min	1st Pctle	99th Pctle	Max
Provided By OPT						
Distance Zone						
A - Less than 0.5 miles	1,043,535	0.33	0	0	1	1
B - Between 0.5 and 1 miles	1,043,535	0.20	0	0	1	1
C - Between 1 and 1.5 miles	1,043,535	0.09	0	0	1	1
D - Greater than 1.5 miles	1,043,535	0.38	0	0	1	1
Missing Distance Zones	41,251					
Transportation Assignment						
Half-fare MetroCard	1,084,786	0.12	0	0	1	1
Full-fare MetroCard	1,084,786	0.34	0	0	1	1
General education school bus	1,084,786	0.07	0	0	1	1
Special education school bus	1,084,786	0.04	0	0	1	1
No assignment	1,084,786	0.43	0	0	1	1
Start Time	959,585	08:07	07:00	07:15	09:00	11:59
End Time	959,585	14:41	12:00	14:00	16:15	17:40
Pick Up Time	76,773	07:22	04:48	05:50	08:15	08:50
Drop Off Time	76,771	15:15	13:36	14:03	17:27	19:45
Calculated						
Transportation Eligibility						
Not eligible	1,043,535	0.20	0	0	1	1
Half-fare MetroCard	1,043,535	0.33	0	0	1	1
Full-fare or school bus	1,043,535	0.20	0	0	1	1
Full-fare	1,043,535	0.27	0	0	1	1
Transportation Exception	1,084,786	0.09	0	0	1	1
Walk time between home and school (minutes)	1,025,464	38.41	0	1.09	267.68	631.83
Drive time between home and school (minutes)	1,025,463	4.88	0	0.29	24.42	66.24
Walk distance between home and school (miles)	1,025,443	2.01	0.01	0.08	14.1	36.33
Drive distance between home and school (minutes)	1,025,466	2.27	0	0.1	14.77	41.82
Walk time between home and bus stop (minutes)	113,369	10.1	0	0.01	211.28	588.75
Walk distance between home and bus stop (miles)	113,360	0.54	0	0.01	11.06	31.6
Time between morning pick up and school start time (minutes)	74,039	38.5	0	5	125	230

Variable Name	Observations	Mean	Min	1st Pctle	99th Pctle	Max
Time from morning pick-up to last stop (minutes)	74,150	13.42	0	0	75	182
Time from last stop to drop-off (minutes)	73,637	3.58	0.04	0.98	12.91	49.47
Time from pick-up to drop-off (minutes)	73,732	17.03	0.04	1.28	79.58	184.17
Time from drop off to start time (minutes)	72,921	21.86	0.21	2.03	88.82	207.73
Time between end of school and afternoon drop off (minutes)	73,155	30.36	0.00	2	135	276
Time between morning pick up and afternoon drop off (hours)	73,227	7.89	6.05	6.55	11.25	13.9
Student attends school which offers a bus	1,084,786	0.44	0	0	1	1

Notes: The calculated bus time variables have different sample sizes because of students missing either pick-up or drop-off times.

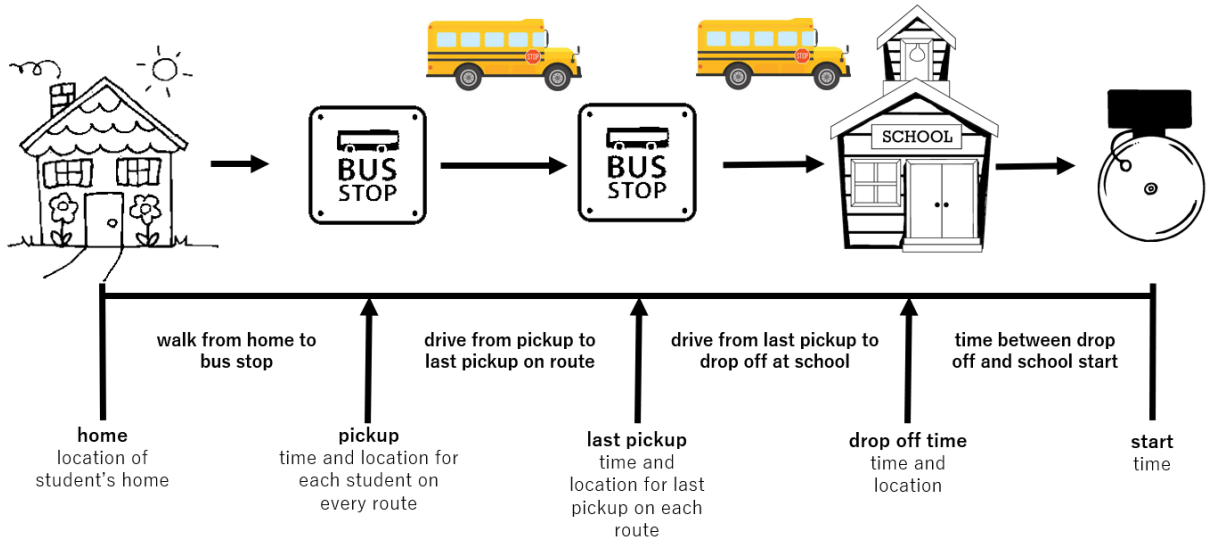
4.3 What are the characteristics of student commutes?

As previously described, we use data on student and school addresses, AM pick-up times, and school start times, to construct various measures of school commutes. On average, students live about 2 miles from school and have walk times of slightly over 30 minutes. However, there is wide variation in these characteristics. For example, while the shortest distance a student travels to school is 0.01 miles, the longest distance is 36.3 miles. Since students tend to live relatively close to school, the drive-time estimate is also short—4.9 minutes—but ranges from less than one minute to 66.2 minutes.

Among bus riders, for whom we have AM pick-up and school start times, we consider four different parts of the commute, illustrated in Figure 1 below: the walk from a student's home to their bus stop, the time from the student's pick-up time to the last pick-up time on the route, the time from last pick-up to school drop-off, and the time between school drop-off and school start times. First, we examine how long/far students travel to reach their assigned bus stop. The average bus rider lives slightly more than a half mile (0.54) or a 10.1 minute walk from their assigned bus stop. Most bus rides are relatively short, or an

average of 13.4 from the student's own stop to the last stop on the bus, and only 3.6 minutes from the last stop to school drop-off. Taken together then, the average amount of time students spend on the bus from their own stop to the school is 17 minutes and the total amount of time they spend from door to door, including the walk to the bus stop, is roughly 27 minutes. Again, there is wide variation in these commute times. While some students have very short bus ride (often because they receive an exception for a hazardous walking condition), others are on the bus for more than two hours.

Figure 1: Legs of the Journey from Home to School



While often ignored in the literature, we also consider the amount of time that bus riders are waiting at school between morning drop-off and school start times. This may be important if students who arrive at school earlier are better able to take advantage of programs such as tutoring or school breakfast. We find that on average, bus riders arrive at school 21.9 minutes before school starts.

Taken together, this suggests that on average, student commutes in NYC are relatively short—around 30 minutes for both students who walk and slightly less for those who take the bus. However, this masks considerable variation as some students experience very short and very long commutes.

5. Special Issues with Transportation Data

5.1 Institutional Context

Knowledge of institutional details and context are key not only for developing research questions and analysis plans focused on pupil transportation, but also for understanding what types of data might be most useful (and available) to request and where to make data requests.

5.1.1 Transportation Routing and Management

The extent to which routing and management are centralized varies from district to district. For example, in NYC, broad rules for transportation and routing take place at the district level, but individual principals decide to offer bus service and are even responsible for assigning students to bus stops. In New Orleans, however, the reverse is true. While the decision to offer bus services is centralized (all schools are required to have a bus), bus routing and management occurs at the school level. The level at which routing and management occurs may shape the amount of variation in bus service and commute times and also has important implications for how to proceed with data requests. In cases where transportation is managed centrally, data can be obtained with a single request, whereas in cases where transportation is managed at the school-level, researchers would likely need to contact each individual school for transportation data. Due to capacity issues, it is also more likely that a centrally managed pupil transportation system would be able to process a data request.

5.1.2 Routing Restrictions

Knowing about a district's routing restrictions such as maximum/minimum number of students on a route, maximum route length, restrictions on pick-up times, etc. is important for conducting data quality checks understanding differences in service provision and identifying empirical strategies. For example, routing restrictions may prevent students from using school buses to attend choice schools far from home, or

students who are eligible for the bus may not receive service because they do not live close to enough other students at their school to create a viable route.

5.1.3 Eligibility and Assignment

Each state or district has their own eligibility criteria and rules for assigning transportation, and it is important to understand these rules, and any potential nuances, as it relates to both data quality checks and identification of empirical methodologies. For example, transportation offices may use non-publicized rules when assigning transportation, such as calculating distance from school based on a student's block rather than precise address. Failing to factor these guidelines into research methodologies may lead to under- or over-estimating the impacts of school transportation and may cause discrepancies when performing data quality checks.

5.1.4 Transportation Exceptions

Almost all districts have exceptions to their transportation policies where otherwise ineligible students are assigned to school buses or other free transportation. Although the types of exceptions offered and application process can vary, some possible reasons for exceptions include injury, safety concerns along the walking route (for example, students having to cross a large boulevard), or living in a homeless shelter. Understanding these exceptions is important for verifying data quality and answering research questions about the equity of transportation policies.

5.1.5 Other Modes of Transportation

While public transportation may be a feasible option in urban locations such as NYC, Chicago, and Washington, D.C., which have robust public transportation systems, other locations may have either no transportation or large transportation deserts, therefore offering students full fare passes for public transportation may not be a feasible option. Understanding the availability of other transportation can

provide insight into what services the district chooses to offer and the equity implications of who does and does not receive transportation.

5.1.6 Charter and Choice Schools

Laws and policies governing the management and financing of charter and choice school transportation vary widely from state to state and can even vary between districts within a state. In NYC, for example, charter schools are treated exactly like public schools in terms of transportation financing and routing for charter school students is managed centrally along with other traditional public-school students. In other districts, however, charter schools may be required to finance their own transportation (which has implications for whether they would offer transportation) and may be responsible for providing the actual bus service (which may have implications for the quality and efficiency of service). This is particularly important for researchers who wish to understand the role of transportation in school choice.

5.2 Best Practices for Using Administrative Data for Research

5.2.1 Data Privacy, Security, and Management

Researchers who already receive student demographic and test score data from a district may need to sign a separate memorandum of understanding (MOU) or non-disclosure agreement (NDA). Most district partners will require that data be stored on a secure server and be accessible only to those working with it. Under federal institutional review board (IRB) rules, data containing scrambled student id numbers and a limited number of variables that can be used to identify students is considered de-identified. In most cases, studies using de-identified data receive exempt status under IRB review. It is also important for researchers to have an internal NDA for use with anyone accessing the data besides the principal investigator that spells out the ways that the data can be used.

There may be concerns from partnering agencies that researchers with access to bus stop times and locations will be able to identify students. This is always a threat when working with student-level administrative data if it contains student addresses. While many bus stops have multiple students, which would lessen concerns about identification, some may only have one or two students assigned. One way to alleviate concerns about student identification is by working with historic data that do not provide information on current students and bus stops. Another option is to suppress any results with fewer than 10 students to avoid any possibility of identification. As an additional protection, researchers can strip any linked data of student addresses, leaving only the x-y coordinates and/or building identifiers for each address. In NYC data, for example, this means removing individual apartment numbers from street addresses and keeping addresses separate from the analytic files.

Finally, once researchers obtain the data, a copy of the raw data should be kept in a folder that can only be accessed by the data manager. A separate, read-only file should be created and stored in a working folder that is accessible to others. This will avoid the original data being over-written in the cleaning or analysis process.

5.2.2 Data Linking

To link transportation data with other records, researchers should ensure that transportation data use the same identification number used by other district offices. In the event the transportation office uses a different id number, the easiest way to link the data is to ask the office that provides the bulk of student-level administrative data to create a new id number and then match the two datasets. To facilitate future matching, this office should keep a record of these new ids along with the algorithm used to create the new id so it can be replicated from year to year. In NYC, researchers received a scrambled identification

number. This number is created each year using the same algorithm so that researchers can follow students over time.

In NYC, we first match transportation data sent by OPT to the demographic data provided from the RPSG by unique student identifier and year and examine the match rate. RPSG provides student-level records from Automate the Schools (ATS) the city's student information system, for all traditional public and charter school students at two times each year, in October and June.[2] The data received from OPT, however, is a snapshot captured at a point in time. As a result, we would expect some students who appear in the data from RPSG but not in the OPT files, and vice versa, but a medium to low match rate between the datasets might indicate larger issues and would warrant further discussion with participating agencies to remedy the problem. For our purposes, we consider students to be enrolled in any academic year where they appear in the data from RPSG and exclude students who are only found in the OPT data. We have a high match rate between the data provided by both offices—about 90% per year. Most of the unmatched records are cases where students are not found in the OPT data, which reflect students who left NYC public schools before October 31 and students in pre-k and alternative programs, which receive different transportation services that are not part of the data requested from OPT.

[2] As of 2019-20, DOE can no longer provide data on charter school students.

References

Blagg, K., Rosenboom, V., & Chingos, M. M. (2018). The Extra Mile: Time to School and Student Outcomes in Washington, DC. Research Report. *Urban Institute*.

Cordes, S. A., Leardo, M., Rick, C., & Schwartz, A. (2019). Can school buses drive down (chronic) absenteeism. *Absent from school: Understanding and addressing student absenteeism*, 121-136.

Cordes, S. A., & Schwartz, A. E. (2018). Does Pupil Transportation Close the School Quality Gap? Evidence from New York City. Research Report. Updated. *Urban Institute*.

Edwards, D.S. (2021). Another One Rides the Bus: The Effects of Transportation Eligibility on School Attendance in Urban and Suburban Michigan. Working Paper.

Gottfried, M. A. (2017). Linking getting to school with going to school. *Educational Evaluation and Policy Analysis*, *39*(4), 571-592.

Gottfried, M. A., Ozuna, C. S., & Kirksey, J. J. (2021). Exploring school bus ridership and absenteeism in rural communities. *Early Childhood Research Quarterly*, *56*, 236-247.

Pogodzinski, B., Lenhoff, S. W., Cook, W., & Singer, J. (2021). School Transit and Accessing Public School in Detroit. *Education and Urban Society*, 00131245211027369.

Stein, M. L., Burdick-Will, J., & Grigg, J. (2021). A choice too far: Transit difficulty and early high school transfer. *Educational Researcher*, *50*(3), 137-144.

Stein, M. L., & Grigg, J. A. (2019). Missing bus, missing school: Establishing the relationship between public transit use and student absenteeism. *American Educational Research Journal*, *56*(5), 1834-1860.

Trajkovski, S., Zabel, J., & Schwartz, A. E. (2021). Do school buses make school choice work? *Regional Science and Urban Economics*, *86*, 103607.

Appendix

As with any administrative data, it is important to run a series of quality checks once pupil transportation data are received from the appropriate agency. Since these data are collected for administrative, rather than research purposes, there may be errors in data entry or compilation that, if unaddressed, could have consequences for researchers' analyses. If there are any discrepancies while performing data quality checks, contact the office or department providing the data to determine possible causes and appropriate solutions. Below, we describe a series of data checks that we perform on the NYC Pupil Transportation Data.

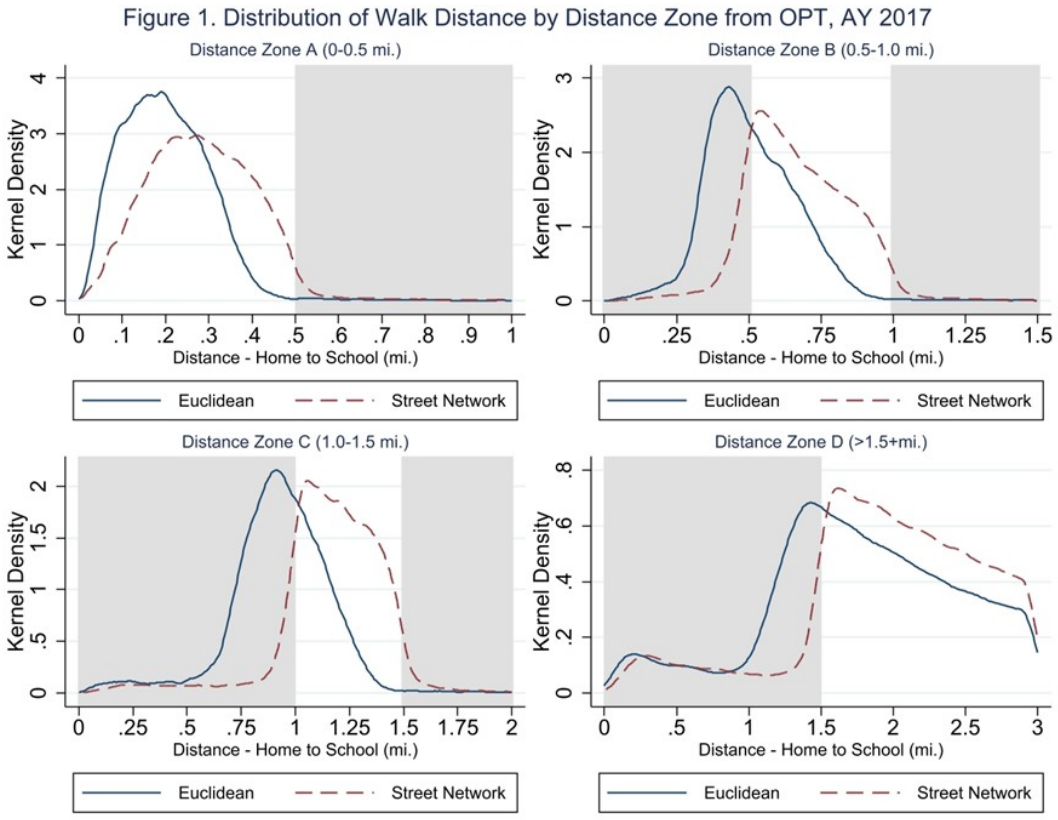
1.1 Bus Eligibility and Assignment

OPT determines eligibility for the bus using a combination of the student's grade and walking distance from home to school. Our data from OPT provide the walk distance category (A, B, C, D) assigned by OPT and the grade for each student, while our data from RPSG provide each student's grade and building of residence. As part of our data checks, we compare the grade reported by OPT to the grade reported from RPSG based on ATS. The match between the two sources is very high (over 99 percent).

We also compare the walk distance category from OPT with the walk distances we calculate based on students' home and school addresses. Initially, we used Euclidean distances to calculate walk distances between students' home and school addresses, but this led us to underestimate walk distances, particularly for those students who live close to school. For example, approximately 30.95% of students' calculated walk distances did not align with distance zones as reported by OPT in 2017. Using street network distances increased our match rates substantially from 69.05% to 83.72% (Figure 1). Most of the improvement was for students who lived between 0.5 and 1.5 miles from school, suggesting that while Euclidean distance and street-network distances are comparable at larger distances, when

calculating distances for students who live closer to school, street-network calculations are preferable. The main trade-off between the two is computational, as street-network distances take more time and computing power to calculate. Thus, in cases where children tend to live far from school and computing power is a concern, Euclidean distances may be preferable.

Figure A1. Distribution of Walk Distance by Distance Zone from OPT, AY 2017



While the street network distances aligned closer with the distance zones reported by OPT, it was more difficult to replicate the OPT assigned distance zones using the data. In order to resolve this remaining discrepancy we contacted OPT and learned that they create the distance categories (A, B, C, and D)

based on *block face* rather than *building* of residence. OPT assigns all students who have the same block face to the same distance category, which is based on the furthest point on the block. For example, if the furthest point on the block is 1 mile from school, all students living on that block would be assigned distance zone B, even if they live closer than 1 mile. This rule was designed to prevent situations where one student receives transportation while another student on the same block does not.

As our final check on eligibility and assignment, we also examine whether transportation assignment aligns with eligibility as defined using OPT-provided distance categories.[3] Only 0.22% of observations in grades K-12 are assigned transportation for which they are not eligible and for which they do not have an exception. Over half of these students were in grades 7 and 8 in Staten Island. Again, through consultation with OPT, we learned that in Staten Island, general education bus eligibility is extended through 8th grade.

1.2 Bus Pick-up and School Start Times

We examine the school bus pick-up times and school start times for general education (GE) students to make sure these were reasonable (i.e., that students are not recorded as being picked up 3:00 am or that schools are not recorded as starting at midnight). Between 2011 and 2017, there are 1,616 observations where students on the GE bus are picked up at their bus stop before 6:00 am (0.34% of GE bus riders) and only 440 observations where students are picked up at their bus stop before 5:30am (0.09%). All school start times are 7:00 am or later.

[3] We exclude students with transportation exceptions from this check. Students with an exception may receive transportation even if they are not eligible based on grade and distance.

We also check to see if bus stop pick-up times occur before school start times. Only a small number of students are reported as being picked up after school starts (1,815 observations from AY 2011-AY 2017, 0.38% of GE bus students), which we drop from our analysis.

1.3 Bus Stops

OPT provides the bus stop coordinates for each student receiving a GE bus. We check the unique number of bus stop coordinates in each year. By doing so, we were able to determine that the bus stop locations in our initial 2012 data pull were incorrect, as all GE bus students were assigned to the same location.

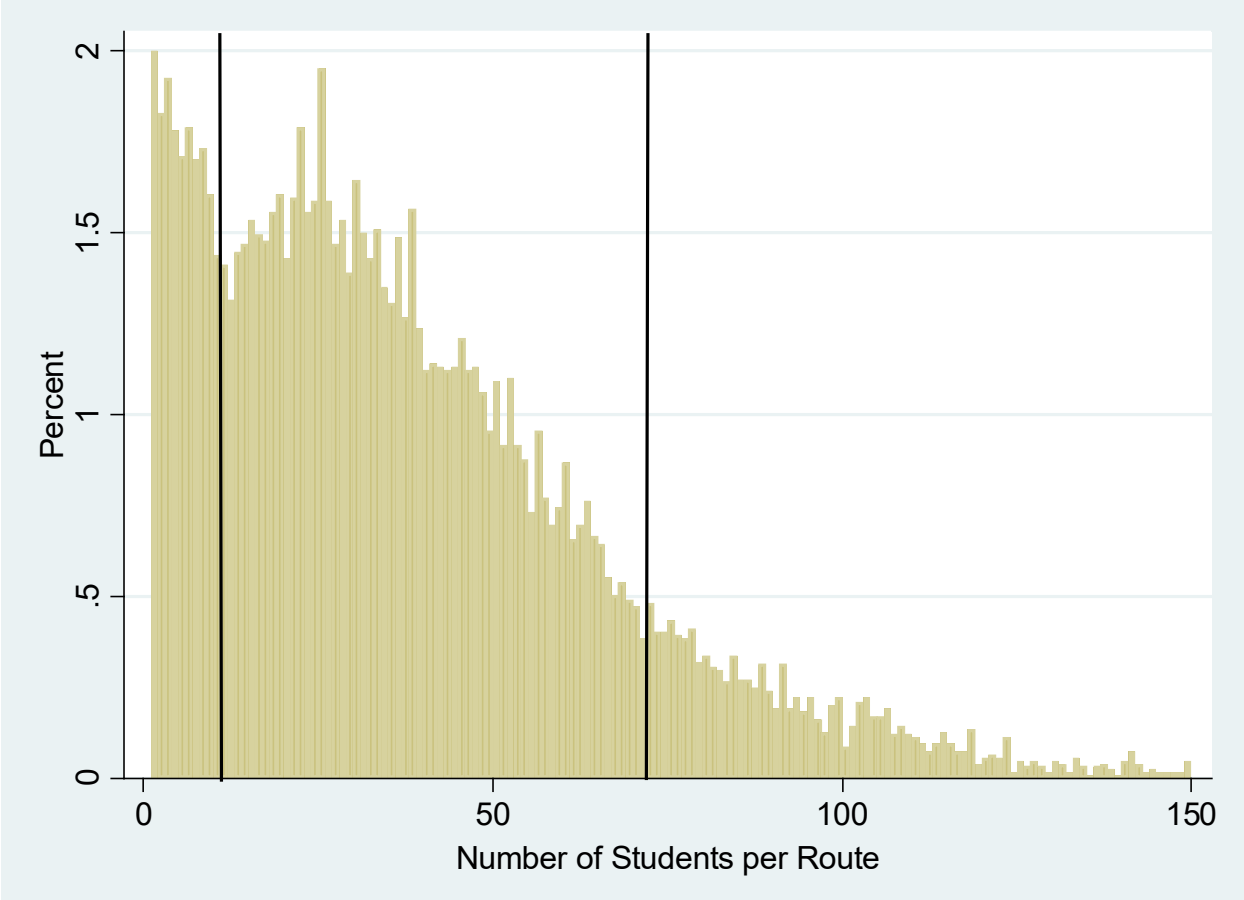
1.4 Bus Routes

NYC has several guidelines to help establish GE bus routes. These include a minimum number of students required to *establish* a bus route (11 students), a maximum distance a bus route can travel (5 miles), and a rule that bus routes for traditional public-school students cannot cross community school boundaries and bus routes for charter school students cannot cross borough boundaries. We use these guidelines to conduct further data checks.

First, OPT guidelines require a minimum of 11 students to establish a general education bus route, although routes may be maintained with fewer than 11 students. Based on this guideline, we examined the number of students on each route and found that 82.54% of morning GE bus routes serve at least 11 students (Figure A2).

We also check to make sure that bus routes serve a plausible number of students, as a standard school bus can fit a maximum of 72 students. A small number of morning general education bus routes have more than 72 students on them (4.34%).

Figure 2. Distribution of number of students on AM general education bus routes, AY 2011-2017



Lastly, we check that routes for traditional public-school students do not cross community school district boundaries. We then check that routes for charter school students do not cross borough boundaries—this is true for 98 of routes serving charter school students.