

Alternate Data File

Study Information:

Title: Impacts of Home Visits on Students in District of Columbia Public Schools

URL: <https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=4597>

Abstract: This study examined the impacts of structured relationship-building teacher home visits conducted in first to fifth grade as part of a family engagement program operating in the District of Columbia Public Schools. Using a matched comparison group research design, the study measured the impacts of the home visits on student disciplinary incidents and attendance. The study found:

- **Home visits prior to the start of the school year reduced the likelihood of a student having a disciplinary incident in that school year.** During the school year following a home visit, 9.27 percent of visited students had a disciplinary incident compared to 12.22 percent of non-visited comparison students.
- **Home visits were associated with a small improvement in student attendance, on average.** Attendance rates averaged 95.28 percent for students who received home visits and 94.93 percent for comparison students who did not.

Authors: Allison McKie, Jeffrey Terziev, and Brian Gill

Principal Investigator's Contact Information:

Allison McKie

amckie@mathematica-mpr.com

(202) 484-4681

Description of the Sample:

Grade levels: 1–5

Any specific subgroups: None.

Sample size: 3,996 student-year observations (1,984 in the home visit group and 2,012 in the comparison group)

Data Information:

Data source #1: Data_Main.dta

School years: 2014/15, 2015/16, and 2016/17

Table 1. Variables Name and Definitions for Data_Main.dta

Variable Name	Variable Definition
stuid	Student identifier
year	Year of spring semester
prioryrdata	Stuid is in attendance file in immediate prior year
prioryrattok	Stuid has attendance data in immediate prior year that is OK to use (i.e., no conflicts or ambiguity in the prior-year attendance data)
f_grade	Grade in stuid-year record with latest admission/withdrawal dates (string)
gradenum	Grade in latest attendance record for a given stuid-year (numeric)
grade45	Indicator for being in grade 4 or 5
p_f_grade	f_grade from immediate prior year
f_schlcode	School code in stuid-year rec with latest admission/withdrawal dates
p_f_schlcode	f_schlcode from immediate prior year

Variable Name	Variable Definition
f schlname	School name in stuid-year rec with latest admission/withdrawal dates
p f schlname	f schlname from immediate prior year
attrate	Stuid-year attendance rate using total (=excused+unexcused) absences
p attrate	attrate from immediate prior year
hvnum	Number of home visits this school year
hv1date	Date of 1st home visit in a given school year
hv1time	Timing of 1st home visit in a given school year
hvprog	Home visit program (0=None, 1=FEP, or 2=FEC/FECMasters)
fep	FEP indicator
hasinc	Stuid had a disciplinary incident during this school year
p daysinc	daysinc from immediate prior year
hasinc t3to5	stuid had a tier 3, 4, or 5 disc. incident this schl year
incldate	Date of 1st disciplinary incident
incltierhi	Highest tier incident stuid had on 1st incident date
male	Indicator for male
ell	Indicator for being an English learner
p ell	Indicator for being an English learner in immediate prior year
frpl	Indicator for receiving free or reduced-price lunch
p frpl	Indicator for receiving free lunch in immediate prior year
hispanic	Indicator for Hispanic
blacknh	Indicator for Black, non-Hispanic
whitenh	Indicator for White, non-Hispanic
othernh	Indicator for non-White, non-Black, non-Hispanic
sped	Indicator for receiving IDEA services
p sped	Indicator for receiving IDEA services in immediate prior year
overage	Indicator for being over age for grade
p overage	Indicator for being over age for grade in immediate prior year
p_cbg_pctlths	% of Census block group with < HS diploma/GED in immediate prior year
zela	ELA z-score (standardized by year-grade among all full academic year students with valid ELA scores)
p zela	zela from immediate prior year
zmath	Math z-score (standardized by year-grade among all full academic year students with valid math scores)
p zmath	zmath from immediate prior year
p mn attrate	School-grade-year mean of attrate for same school s in grade g-1 in year y-1
p mn daysinc	School-grade-year mean of daysinc for same school s in grade g-1 in year y-1
p mn male	School-grade-year mean of male for same school s in grade g-1 in year y-1
p mn frpl	School-grade-year mean of frpl for same school s in grade g-1 in year y-1
p mn hispanic	School-grade-year mean of hispanic for same school s in grade g-1 in year y-1
p mn blacknh	School-grade-year mean of blacknh for same school s in grade g-1 in year y-1
p mn whitenh	School-grade-year mean of whitenh for same school s in grade g-1 in year y-1
p mn othernh	School-grade-year mean of othernh for same school s in grade g-1 in year y-1
p mn ell	School-grade-year mean of ell for same school s in grade g-1 in year y-1
p mn sped	School-grade-year mean of sped for same school s in grade g-1 in year y-1
p mn overage	School-grade-year mean of overage for same school s in grade g-1 in year y-1
p mn cbg_pctlths	School-grade-year mean of cbg_pctlths for same school s in grade g-1 in year y-1

Variable Name	Variable Definition
p mn zela	School-grade-year mean of zela for same school s in grade g-1 in year y-1
p mn zmath	School-grade-year mean of zmath for same school s in grade g-1 in year y-1
extyr	School had an extended year this year
yrstart	1st day of school year
cat4raceeth	4-category race/ethnicity
hvltime misscat	Timing of 1st home visit this school year, with missing date category
hvlb4yrstart	Treatment student's 1st visit this year happened before 1st day of school year
schl gr yr	School-grade-year identifier
miss p ell	Indicator for missing p ell
miss p frpl	Indicator for missing p frpl
miss p sped	Indicator for missing p sped
miss p overage	Indicator for missing p overage
miss p cbg_pctlths	Indicator for missing p cbg_pctlths
miss p attrate	Indicator for missing p attrate
miss p daysinc	Indicator for missing p daysinc
miss p zela	Indicator for missing p zela
miss p zmath	Indicator for missing p zmath
mi miss	Incomplete observation
1 p attrate	p attrate from imputation 1
2 p attrate	p attrate from imputation 2
3 p attrate	p attrate from imputation 3
4 p attrate	p attrate from imputation 4
5 p attrate	p attrate from imputation 5
1 p daysinc	p daysinc from imputation 1
2 p daysinc	p daysinc from imputation 2
3 p daysinc	p daysinc from imputation 3
4 p daysinc	p daysinc from imputation 4
5 p daysinc	p daysinc from imputation 5
1 p ell	p ell from imputation 1
2 p ell	p ell from imputation 2
3 p ell	p ell from imputation 3
4 p ell	p ell from imputation 4
5 p ell	p ell from imputation 5
1 p sped	p sped from imputation 1
2 p sped	p sped from imputation 2
3 p sped	p sped from imputation 3
4 p sped	p sped from imputation 4
5 p sped	p sped from imputation 5
1 p frpl	p frpl from imputation 1
2 p frpl	p frpl from imputation 2
3 p frpl	p frpl from imputation 3
4 p frpl	p frpl from imputation 4
5 p frpl	p frpl from imputation 5
1 p overage	p overage from imputation 1
2 p overage	p overage from imputation 2
3 p overage	p overage from imputation 3

Variable Name	Variable Definition
4 p overage	p overage from imputation 4
5 p overage	p overage from imputation 5
1 p cbg pctlths	p cbg pctlths from imputation 1
2 p cbg pctlths	p cbg pctlths from imputation 2
3 p cbg pctlths	p cbg pctlths from imputation 3
4 p cbg pctlths	p cbg pctlths from imputation 4
5 p cbg pctlths	p cbg pctlths from imputation 5
1 p zela	p zela from imputation 1
2 p zela	p zela from imputation 2
3 p zela	p zela from imputation 3
4 p zela	p zela from imputation 4
5 p zela	p zela from imputation 5
1 p zmath	p zmath from imputation 1
2 p zmath	p zmath from imputation 2
3 p zmath	p zmath from imputation 3
4 p zmath	p zmath from imputation 4
5 p zmath	p zmath from imputation 5
mnps_fep	Mean FEP propensity score (across imputations, by year-gradegroup [grades 1-3 or grades 4-5])
wgt nnlnor	Observation weight (nearest neighbor 1, no replacement)
id nnlnor	psmatch2 ID for this observation (nearest neighbor 1, no replacement)
tmatch nnlnor	psmatch2 ID for the treatment that defines the matched set (nn1, no repl.)
matchct nnlnor	Number of records in the matched set (nearest neighbor 1, no replacement)

Data source #2: Data IncidentBeforeVisit_FEC.dta

School years: 2014/15, 2015/16, and 2016/17

Table 2. Variables Name and Definitions for Data IncidentBeforeVisit_FEC.dta

Variable Name	Variable Definition
stuid	Student identifier
year	Year of spring semester
prioryrdata	Stuid is in attendance file in immediate prior year
prioryrattok	Stuid has attendance data in immediate prior year that is OK to use (i.e., no conflicts or ambiguity in the prior-year attendance data)
f_grade	Grade in stuid-year record with latest admission/withdrawal dates (string)
gradenum	Grade in latest attendance record for a given stuid-year (numeric)
grade45	Indicator for being in grade 4 or 5
p_f_grade	f_grade from immediate prior year
f_schlcode	School code in stuid-year rec with latest admission/withdrawal dates
p_f_schlcode	f_schlcode from immediate prior year
f_schlname	School name in stuid-year rec with latest admission/withdrawal dates
p_f_schlname	f_schlname from immediate prior year
attrate	Stuid-year attendance rate using total (=excused+unexcused) absences
p_attrate	attrate from immediate prior year
hvnum	Number of home visits this school year
hv1date	Date of treatment student's 1st visit this school year

Variable Name	Variable Definition
hv1time	Timing of treatment student's 1st visit this school year
hvprog	Home visit program (0=None, 1=FEP, or 2=FEC/FECMasters)
fec	FEC (including FEC Masters) indicator
hasinc	Stuid had a disciplinary incident during this school year
p_daysinc	daysinc from immediate prior year
hasinc_t3to5	stuid had a tier 3, 4, or 5 disc. incident this schl year
incldate	Date of 1st disciplinary incident
incltierhi	Highest tier incident stuid had on 1st incident date
male	Indicator for male
ell	Indicator for being an English learner
p_ell	Indicator for being an English learner in immediate prior year
frpl	Indicator for receiving free or reduced-price lunch
p_frpl	Indicator for receiving free lunch in immediate prior year
hispanic	Indicator for Hispanic
blacknh	Indicator for Black, non-Hispanic
whitenh	Indicator for White, non-Hispanic
othernh	Indicator for non-White, non-Black, non-Hispanic
sped	Indicator for receiving IDEA services
p_sped	Indicator for receiving IDEA services in immediate prior year
overage	Indicator for being over age for grade
p_overage	Indicator for being over age for grade in immediate prior year
p_cbg_pctlths	% of Census block group with < HS diploma/GED in immediate prior year
zela	ELA z-score (standardized by year-grade among all full academic year students with valid ELA scores)
p_zela	zela from immediate prior year
zmath	Math z-score (standardized by year-grade among all full academic year students with valid math scores)
p_zmath	zmath from immediate prior year
p_mn_attrate	School-grade-year mean of attrate for same school s in grade g-1 in year y-1
p_mn_daysinc	School-grade-year mean of daysinc for same school s in grade g-1 in year y-1
p_mn_male	School-grade-year mean of male for same school s in grade g-1 in year y-1
p_mn_frpl	School-grade-year mean of frpl for same school s in grade g-1 in year y-1
p_mn_hispanic	School-grade-year mean of hispanic for same school s in grade g-1 in year y-1
p_mn_blacknh	School-grade-year mean of blacknh for same school s in grade g-1 in year y-1
p_mn_whitenh	School-grade-year mean of whitenh for same school s in grade g-1 in year y-1
p_mn_othernh	School-grade-year mean of othernh for same school s in grade g-1 in year y-1
p_mn_ell	School-grade-year mean of ell for same school s in grade g-1 in year y-1
p_mn_sped	School-grade-year mean of sped for same school s in grade g-1 in year y-1
p_mn_overage	School-grade-year mean of overage for same school s in grade g-1 in year y-1
p_mn_cbg_pctlths	School-grade-year mean of cbg_pctlths for same school s in grade g-1 in year y-1
p_mn_zela	School-grade-year mean of zela for same school s in grade g-1 in year y-1
p_mn_zmath	School-grade-year mean of zmath for same school s in grade g-1 in year y-1
extyr	School had an extended year this year
yrstart	1st day of school year
cat4raceeth	4-category race/ethnicity
hv1time_misscat	Timing of treatment student's 1st visit this school year, w/ missing date category

Variable Name	Variable Definition
hv1b4yrstart	Treatment student's 1st visit this year happened before 1st day of school year
schl gr yr	School-grade-year identifier
miss p ell	Indicator for missing p ell
miss p frpl	Indicator for missing p frpl
miss p sped	Indicator for missing p sped
miss p overage	Indicator for missing p overage
miss p cbg pctlths	Indicator for missing p cbg pctlths
miss p attrate	Indicator for missing p attrate
miss p daysinc	Indicator for missing p daysinc
miss p zela	Indicator for missing p zela
miss p zmath	Indicator for missing p zmath
mi miss	Incomplete observation
1 p attrate	p attrate from imputation 1
2 p attrate	p attrate from imputation 2
3 p attrate	p attrate from imputation 3
4 p attrate	p attrate from imputation 4
5 p attrate	p attrate from imputation 5
1 p daysinc	p daysinc from imputation 1
2 p daysinc	p daysinc from imputation 2
3 p daysinc	p daysinc from imputation 3
4 p daysinc	p daysinc from imputation 4
5 p daysinc	p daysinc from imputation 5
1 p ell	p ell from imputation 1
2 p ell	p ell from imputation 2
3 p ell	p ell from imputation 3
4 p ell	p ell from imputation 4
5 p ell	p ell from imputation 5
1 p sped	p sped from imputation 1
2 p sped	p sped from imputation 2
3 p sped	p sped from imputation 3
4 p sped	p sped from imputation 4
5 p sped	p sped from imputation 5
1 p frpl	p frpl from imputation 1
2 p frpl	p frpl from imputation 2
3 p frpl	p frpl from imputation 3
4 p frpl	p frpl from imputation 4
5 p frpl	p frpl from imputation 5
1 p overage	p overage from imputation 1
2 p overage	p overage from imputation 2
3 p overage	p overage from imputation 3
4 p overage	p overage from imputation 4
5 p overage	p overage from imputation 5
1 p cbg pctlths	p cbg pctlths from imputation 1
2 p cbg pctlths	p cbg pctlths from imputation 2
3 p cbg pctlths	p cbg pctlths from imputation 3
4 p cbg pctlths	p cbg pctlths from imputation 4

Variable Name	Variable Definition
5 p cbg pctlths	p cbg pctlths from imputation 5
1 p zela	p zela from imputation 1
2 p zela	p zela from imputation 2
3 p zela	p zela from imputation 3
4 p zela	p zela from imputation 4
5 p zela	p zela from imputation 5
1 p zmath	p zmath from imputation 1
2 p zmath	p zmath from imputation 2
3 p zmath	p zmath from imputation 3
4 p zmath	p zmath from imputation 4
5 p zmath	p zmath from imputation 5
mnps_fec	Mean FEC propensity score (across imputations, by year-gradedgroup [grades 1-3 or grades 4-5])
wgt_nnlnor	Observation weight (nearest neighbor 1, no replacement)
id_nnlnor	psmatch2 ID for this observation (nearest neighbor 1, no replacement)
tmatch_nnlnor	psmatch2 ID for the treatment that defines the matched set (nn1, no repl.)
matchct_nnlnor	Number of records in the matched set (nearest neighbor 1, no replacement)
incb4hv1	Had a disciplinary incident before treatment student's 1st visit this school year

Data source #3: Data IncidentBeforeVisit_FEP.dta

School years: 2014/15, 2015/16, and 2016/17

Table 3. Variables Name and Definitions for Data IncidentBeforeVisit_FEP.dta

Variable Name	Variable Definition
stuid	Student identifier
year	Year of spring semester
prioryrdata	Stuid is in attendance file in immediate prior year
prioryrattok	Stuid has attendance data in immediate prior year that is OK to use (i.e., no conflicts or ambiguity in the prior-year attendance data)
f_grade	Grade in stuid-year record with latest admission/withdrawal dates (string)
gradenum	Grade in latest attendance record for a given stuid-year (numeric)
grade45	Indicator for being in grade 4 or 5
p_f_grade	f_grade from immediate prior year
f_schlcode	School code in stuid-year rec with latest admission/withdrawal dates
p_f_schlcode	f_schlcode from immediate prior year
f_schlname	School name in stuid-year rec with latest admission/withdrawal dates
p_f_schlname	f_schlname from immediate prior year
attrate	Stuid-year attendance rate using total (=excused+unexcused) absences
p_attrate	attrate from immediate prior year
hvnum	Number of home visits this school year
hv1date	Date of treatment student's 1st visit this school year
hv1time	Timing of treatment student's 1st visit this school year
hvprog	Home visit program (0=None, 1=FEP, or 2=FEC/FECMasters)
fep	FEP indicator
hasinc	Stuid had a disciplinary incident during this school year
p_daysinc	daysinc from immediate prior year

Variable Name	Variable Definition
hasinc_t3to5	stuid had a tier 3, 4, or 5 disc. incident this schl year
incldate	Date of 1st disciplinary incident
incltierhi	Highest tier incident stuid had on 1st incident date
male	Indicator for male
ell	Indicator for being an English learner
p_ell	Indicator for being an English learner in immediate prior year
frpl	Indicator for receiving free or reduced-price lunch
p_frpl	Indicator for receiving free lunch in immediate prior year
hispanic	Indicator for Hispanic
blacknh	Indicator for Black, non-Hispanic
whitenh	Indicator for White, non-Hispanic
othernh	Indicator for non-White, non-Black, non-Hispanic
sped	Indicator for receiving IDEA services
p_sped	Indicator for receiving IDEA services in immediate prior year
overage	Indicator for being over age for grade
p_overage	Indicator for being over age for grade in immediate prior year
p_cbg_pctlths	% of Census block group with < HS diploma/GED in immediate prior year
zela	ELA z-score (standardized by year-grade among all full academic year students with valid ELA scores)
p_zela	zela from immediate prior year
zmath	Math z-score (standardized by year-grade among all full academic year students with valid math scores)
p_zmath	zmath from immediate prior year
p_mn_attrate	School-grade-year mean of attrate for same school s in grade g-1 in year y-1
p_mn_daysinc	School-grade-year mean of daysinc for same school s in grade g-1 in year y-1
p_mn_male	School-grade-year mean of male for same school s in grade g-1 in year y-1
p_mn_frpl	School-grade-year mean of frpl for same school s in grade g-1 in year y-1
p_mn_hispanic	School-grade-year mean of hispanic for same school s in grade g-1 in year y-1
p_mn_blacknh	School-grade-year mean of blacknh for same school s in grade g-1 in year y-1
p_mn_whitenh	School-grade-year mean of whitenh for same school s in grade g-1 in year y-1
p_mn_othernh	School-grade-year mean of othernh for same school s in grade g-1 in year y-1
p_mn_ell	School-grade-year mean of ell for same school s in grade g-1 in year y-1
p_mn_sped	School-grade-year mean of sped for same school s in grade g-1 in year y-1
p_mn_overage	School-grade-year mean of overage for same school s in grade g-1 in year y-1
p_mn_cbg_pctlths	School-grade-year mean of cbg_pctlths for same school s in grade g-1 in year y-1
p_mn_zela	School-grade-year mean of zela for same school s in grade g-1 in year y-1
p_mn_zmath	School-grade-year mean of zmath for same school s in grade g-1 in year y-1
extyr	School had an extended year this year
yrstart	1st day of school year
cat4raceeth	4-category race/ethnicity
hv1time_misscat	Timing of treatment student's 1st visit this school year, w/ missing date category
hv1b4yrstart	Treatment student's 1st visit this year happened before 1st day of school year
schl_gr_yr	School-grade-year identifier
miss_p_ell	Indicator for missing p_ell
miss_p_frpl	Indicator for missing p_frpl
miss_p_sped	Indicator for missing p_sped

Variable Name	Variable Definition
miss p_ouverage	Indicator for missing p_ouverage
miss p_cbg_pctlths	Indicator for missing p_cbg_pctlths
miss p_attrate	Indicator for missing p_attrate
miss p_daysinc	Indicator for missing p_daysinc
miss p_zela	Indicator for missing p_zela
miss p_zmath	Indicator for missing p_zmath
mi_miss	Incomplete observation
1 p_attrate	p_attrate from imputation 1
2 p_attrate	p_attrate from imputation 2
3 p_attrate	p_attrate from imputation 3
4 p_attrate	p_attrate from imputation 4
5 p_attrate	p_attrate from imputation 5
1 p_daysinc	p_daysinc from imputation 1
2 p_daysinc	p_daysinc from imputation 2
3 p_daysinc	p_daysinc from imputation 3
4 p_daysinc	p_daysinc from imputation 4
5 p_daysinc	p_daysinc from imputation 5
1 p_ell	p_ell from imputation 1
2 p_ell	p_ell from imputation 2
3 p_ell	p_ell from imputation 3
4 p_ell	p_ell from imputation 4
5 p_ell	p_ell from imputation 5
1 p_sped	p_sped from imputation 1
2 p_sped	p_sped from imputation 2
3 p_sped	p_sped from imputation 3
4 p_sped	p_sped from imputation 4
5 p_sped	p_sped from imputation 5
1 p_frpl	p_frpl from imputation 1
2 p_frpl	p_frpl from imputation 2
3 p_frpl	p_frpl from imputation 3
4 p_frpl	p_frpl from imputation 4
5 p_frpl	p_frpl from imputation 5
1 p_ouverage	p_ouverage from imputation 1
2 p_ouverage	p_ouverage from imputation 2
3 p_ouverage	p_ouverage from imputation 3
4 p_ouverage	p_ouverage from imputation 4
5 p_ouverage	p_ouverage from imputation 5
1 p_cbg_pctlths	p_cbg_pctlths from imputation 1
2 p_cbg_pctlths	p_cbg_pctlths from imputation 2
3 p_cbg_pctlths	p_cbg_pctlths from imputation 3
4 p_cbg_pctlths	p_cbg_pctlths from imputation 4
5 p_cbg_pctlths	p_cbg_pctlths from imputation 5
1 p_zela	p_zela from imputation 1
2 p_zela	p_zela from imputation 2
3 p_zela	p_zela from imputation 3
4 p_zela	p_zela from imputation 4

Variable Name	Variable Definition
5 p zela	p zela from imputation 5
1 p zmath	p zmath from imputation 1
2 p zmath	p zmath from imputation 2
3 p zmath	p zmath from imputation 3
4 p zmath	p zmath from imputation 4
5 p zmath	p zmath from imputation 5
mnps_fep	Mean FEP propensity score (across imputations, by year-gradegroup [grades 1-3 or grades 4-5])
wgt_nn1nor	Observation weight (nearest neighbor 1, no replacement)
id_nn1nor	psmatch2 ID for this observation (nearest neighbor 1, no replacement)
tmatch_nn1nor	psmatch2 ID for the treatment that defines the matched set (nn1, no repl.)
matchct_nn1nor	Number of records in the matched set (nearest neighbor 1, no replacement)
incb4hv1	Had a disciplinary incident before treatment student's 1st visit this school year

Data source #4: Data_DemogsForAllDCPS.dta

School years: 2014/15, 2015/16, and 2016/17

Table 4. Variables Name and Definitions for Data_DemogsForAllDCPS.dta

Variable Name	Variable Definition
stuid	Student identifier
year	Year of spring semester
male	Indicator for male
ell	Indicator for being an English learner
frpl	Indicator for receiving free lunch
hispanic	Indicator for Hispanic
blacknh	Indicator for Black, non-Hispanic
whitenh	Indicator for White, non-Hispanic
othernh	Indicator for non-White, non-Black, non-Hispanic
sped	Indicator for receiving IDEA services
overage	Indicator for being over age for grade

Data Source #5: Data_PgmDistributions.dta

School years: 2013/14, 2014/15, 2015/16, and 2016/17

Table 5. Variables Name and Definitions for Data_PgmDistributions.dta

Variable Name	Variable Definition
stuid	Student identifier
year	Year of spring semester
hvprog	Home visit program (0=None, 1=FEP, or 2=FEC/FECMasters)
hv1b4yrstart	1st visit this year happened before 1st day of school year

Analysis Information:

- **Data analyzed.** All data were obtained for the 2012/13–2016/17 school years, unless otherwise noted. The Flamboyan Foundation owns the FEP data and provided the data to the study team via DCPS. All other data are owned and provided by DCPS, except for the census block group characteristics, which are publicly available on the U.S. Census Bureau website (<https://data.census.gov>). The study team used the following data:
 - **Home visit data:** Name, grade, and school of students who received home visits; names of teachers who provided those visits; and the dates of the visits. We obtained these data for the 2013/14 – 2016/17 school years. (We requested data for the 2012/13 school year but DCPS indicated those data are no longer available. FEP visit dates are not available for the 2013/14 school year.)
 - **Student attendance and enrollment status:** Student-level grade, school, admission date, withdrawal date, membership days (i.e., number of days enrolled), excused absences, and unexcused absences.
 - **Student office referrals:** Date and tier of each disciplinary incident (see table B2 in the report appendices at <https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=4597> for a description of the tiers).
 - **Student achievement:** Student-level English language arts and math scale scores from the DC CAS for the 2012/13 and 2013/14 school years and from PARCC assessments for the 2014/15, 2015/16, and 2016/17 school years. To express scores in a common unit, we used means and standard deviations of DC CAS and PARCC scale scores for the DCPS test-taking population to convert scale scores into *z*-scores (see table B2 in the report appendices for additional information on *z*-scores).
 - **Student demographic data and census block group characteristics:**
 - **DCPS data:** Student-level race/ethnicity, gender, age, free or reduced-price lunch status, special education status, English learner status, and home address.
 - **Census block group characteristics:** Five-year averages in each U.S. census block group of the percentage of individuals age 25 and older by highest educational attainment. The study team attempted to match each student in the sample to his or her census block group characteristics using student home address data.
 - **Class rosters and teacher data.** The study team did not include teacher data in the analysis after we found that we could not reliably link students to their classroom teachers. We did, however, use student, teacher, and school identifiers from class roster and teacher data in cleaning the home visit data, as discussed below.
- **Cleaning home visit data.** Cleaning the home visit data proceeded in two stages.
 - **Stage 1.** The first stage of the cleaning process aimed to produce a dataset consisting of one record for each student visit in each school year, with that student-year record ideally including the student’s name, school, grade, visit program, visit date, and the names of each of the two teachers who conducted the visit. For example, the raw visit data might include two separate records for the same student visit, with each record listing only one of the two visiting teachers; or a student’s name might vary slightly across records. The study team sought to standardize key data fields and consolidate records for a given student visit. The raw records also varied in their completeness. To the extent possible, we used available data to fill in missing information. For instance, if a record omitted the student’s school, we could often identify the school based on the visit date and the names of the visiting teachers using DCPS teacher data.
 - **Stage 2.** The second cleaning stage sought to determine the unique numeric DCPS student identifier (renamed “stuid” in our datasets) for each grade 1-5 student in the home visit

data each school year, thus enabling the merging of home visit and DCPS administrative data. In Stata, we ran a series of match merges between the home visit dataset created in the preceding stage and DCPS attendance files and used the record linkage command *relink2* to identify stuids for grade-eligible students. We also used DCPS class roster and teacher data to investigate potential matches as needed. Potential matches that remained ambiguous were dropped from the study sample to avoid misclassifying a student's treatment status. As an illustrative example using fictitious names and stuids, if available data indicated that the student listed in the 2014/15 home visit data as Jayne Smyth could be either stuid 11111 with name Jane Smyth or stuid 22222 with name Jayne Smith, we dropped stuids 11111 and 22222 from the sample in 2014/15.

- **Cleaning other data.** Cleaning the non-home-visit data was relatively straightforward. We note below key decisions made or processes used for each type of data.
 - **Student attendance and enrollment status.** Because DCPS advised us to obtain key enrollment information such as grade and school from the attendance files, we used the attendance files to define our sample. We excluded from the analysis any stuid-years that did not appear in the attendance files. We also excluded a few stuid-years that had conflicting or ambiguous attendance data (e.g., a stuid-year with multiple attendance records that indicated different enrollment and absence data for the same school year).
 - **Student disciplinary incidents obtained from office referral files.** We created disciplinary incident variables using primary records (i.e., records for which DCPS variable “isprimary” equaled 1).
 - **Student achievement.** We attempted to mimic aspects of the DC assessment policy by retaining only valid DC CAS or PARCC test scores for students who met the full academic year criteria for at least one level (school, local education agency, or state) and took the assessment for the grade in which they were enrolled at the time of testing.
 - **Student demographics and census block group characteristics.**
 - **Time-invariant student demographics.** To minimize missing data and resolve any discrepancies across years in race/ethnicity, gender, and date of birth, we used the most recent information available for each stuid.
 - **Special education.** The demographic files we obtained did not include special education status for the 2012/13 school year and had a relatively large proportion of students missing special education status for the 2013/14 school year. The DC CAS files included special education status for students in grades 2-10 for the 2012/13 school year and for students in grades 3 and higher for the 2013/14 school year. When available, we used special education status from the DC CAS for all students in 2012/13 and for students missing special education data in the 2013/14 demographic file.
 - **Census block group characteristics.** After cleaning and standardizing student home addresses obtained from the DCPS demographic files, we used the [United States Census Bureau's Geocoder](#) to obtain a census block group identifier for each home address. We then used that identifier to merge census block group characteristics (available at <https://www.census.gov/data.html>) to the student files.
- **Creating the final analytic data files.** After cleaning the datasets, we constructed the analytic variables and merged the home visit, disciplinary incident, achievement, demographic, and Census block group data to the attendance data by stuid and year. We then created multiply imputed datasets, estimated propensity scores, and conducted matching. Report appendix B describes our methods. Table 6 below identifies the final analytic dataset used to obtain each set of report findings.

Table 6. Datasets Used to Obtain Report Findings

Dataset	Report Findings
Data_Main.dta	Impacts of FEP summer home visits on student disciplinary incidents (table 2), attendance (table 3), and achievement (table 4)
	Percentages of baseline characteristics data that are imputed (table b3)
	Baseline characteristics of analytic samples for impacts of FEP summer visits on disciplinary incidents and attendance for grades 1–5 (table c1) and achievement for grades 4 and 5 (table c2)
	Characteristics used to assess What Works Clearinghouse baseline equivalence requirements (table c3)
Data_IncidentBeforeVisit_FEC.dta	Likelihood of FEC students having a disciplinary incident by the time of the visited student’s first visit of the school year (table d3)
Data_IncidentBeforeVisit_FEP.dta	Likelihood of FEP students having a disciplinary incident by the time of the visited student’s first visit of the school year (table d3)
Data_DemogsForAllDCPS.dta	Demographic characteristics of FEP-visited students and all DCPS students in grades 1–5 (table b1)
Data_PgmDistributions.dta	Distribution of visited students across programs (FEC and FEP) by school year (table d1)
	Distribution of FEP students by timing of their first visit of the school year (table d4)
	Distribution of FEC students timing of their first visit of the school year (table d5)

Notes: FEC = Family Engagement Collaborative. FEP = Family Engagement Partnership.

Contact information:

To request the study data, please follow the District of Columbia Public Schools’ (DCPS) process for obtaining confidential data found here: <https://dcps.dc.gov/page/conduct-research-or-obtain-confidential-data>. In the request, ask for the data from the Evaluation of the Impact of Structured Relationship-Building Teacher Home Visiting on Student Outcomes, refer to the DCPS folder code Mathematica_Home Visit, and mention that the requested data are from the Memorandum of Agreement (MOA) that was fully executed on June 6, 2018. If you have questions about the process, please email researchrequests@k12.dc.gov.