



Department
for Education

The National Pupil Database

Data Request Application Form

Introduction

National pupil database data

This application form should be used for requesting extracts of data from the National Pupil Database (NPD). Other documents relating to the data request process can be found on the [NPD Online](#) website.

Before making an application, please make sure that you have read all the accompanying information on the [NPD Online](#) website including the [NPD User Guide](#).

This form should be filled in electronically and e-mailed to the NPD Data Request team at NPD.REQUESTS@education.gsi.gov.uk with a completed [Information Security Questionnaire](#).

Who can request access to the data?

The Education (Individual Pupil Information) (Prescribed Persons) (England) Regulations 2009, as amended enable the Department for Education to share individual pupil information from the NPD with named bodies and persons who for the purpose of promoting the education or well-being of children in England are conducting research or analysis, producing statistics, or providing information, advice or guidance. Access is subject to requesters complying with terms and conditions imposed under contractual arrangements and a rigorous approvals process.

Any person requesting access to the data will have to demonstrate that they will comply with all relevant requirements of the Data Protection Act 1998. They will need to demonstrate that they:

- Have appropriate security arrangements in place to process the data;
- Intend to use the data only for the specified purpose;
- Will keep the data only for the specified length of time;
- Will not try to further disclose the data without the prior written approval of the Department for Education.

Applicants will be expected to sign an agreement to confirm this and provide additional information about their security arrangements.

Please keep a copy of all the documents, including completed forms, for your own records.

Linked data

If you are applying for linked data you need to apply using a different form which can be found online [here](#). Linked data is any combination of the following:

- NPD data linked to Individualised Learner Record (ILR) data;
- NPD data linked to Higher Education Statistics Agency (HESA) data;
- NPD data linked to ILR and HESA data; or
- ILR data linked to HESA data.

Processing of NPD data

When you make a request for NPD data it will be considered for approval by the Education Data Division (EDD) with the exception of tier 1 data requests, which will be assessed by the department's Data Management Advisory Panel. The EDD will inform you of the outcome of the decision.

Where it is agreed to provide you with data, you will be provided with a [Data Sharing Agreement](#) and an [Individual Declaration Form](#) to sign and return.

Publication of details of requests

Please note that the Department for Education (DfE) has a policy of publishing details of all requests it receives for NPD data on their website.

DfE will publish a description of the NPD data requested and whether or not it is requested to be linked to ILR data or HESA data, the name of the requesting organisation, the intended use of the data and details of the outcome. For the avoidance of doubt, no personal data will be published by DfE.

Application Form

1	Contact details
Name of Applicant	
[REDACTED]	
Name of person / organisation who will be entering into the Data Sharing Agreement ("the Requester")	
[REDACTED] (Stanford University)	
Organisation type (e.g. Academic institution, Research organisation, Media, Individual citizen, etc.)	
Academic Institution	
Position	
Graduate Students	
Department	
Economics	
Correspondence Address	
579 Serra Mall Stanford, CA 94305 United States of America	
Postcode	
94305	
Telephone	
[REDACTED]	
Address at which the data will be processed (if different from above)	
579 Serra Mall Stanford, CA 94305 United States of America	
Postcode	
94305	
Telephone	
[REDACTED]	
If you are a student please specify your supervising tutor and their telephone and e-mail address	
[REDACTED]	

2 Details of DfE Project Sponsor (if applicable)	
Title (Mr, Mrs, Ms, etc.)	
Name	
Job Title	
Division / Team	
Contact Address	
Telephone	
E-mail address	

3 Have you (or your organisation) previously applied to use or access NPD data (Please insert X)	
Yes - state reference number(s), date of application(s) and name of applicant(s)	
No	X

4 Information Commissioner's Office (ICO) Registration	
a	ICO Registration Number or Exemption Ground (It is mandatory to provide a registration number unless you have a valid exemption)
ZA130418	
b	ICO Registration Expiry Date
21 July 2016	

5 Enquiry details	
NPD - Specify which datasets, versions, tiers, academic years and sensitive fields are required along with any specific requirements relating to data linking and coverage in line with the NPD published guidance .	
<p>The following list includes the dataset we would like to request. Required variables from Tier 1 and 2 are explicitly listed. We would like to obtain access to all Tier 3 and 4 variables.</p> <p><u>PLASC Census 01-02 to 13-14 SUM, Tier 2, 4, AY: 2006-07 - 2013-2014 (or latest)</u></p> <p>Sensitive Fields:</p> <ul style="list-style-type: none"> • AgeAtStartOfAcademicYear • EthnicGroupMajor • FSMeligible • EntryDate • LeavingDate 	

- SENprovision
- DistCurrSch
- DistNearSch
- NearSchLAEstab
- LSOA01, LSOA11

KS1 97-98 to 13-14, Tier 4, AY: 2006-07 -2013-2014 (or latest)

KS2 95-96 to 13-14, Tier 4, AY: 2006-07 -2013-2014 (or latest)

Year 7 00-01 to 06-07, Tier 4, AY: 2006-07 2013-2014 (or latest)

KS3 97-98 to 12-13, Tier 4, AY: 2006-07 -2013-2014 (or latest)

KS4 01-02 to 13-14 (2013 methodology), Tier 4, AY: 2006-07 2013-2014 (or latest)

KS4 13-14 (2014 methodology), Tier 4, AY: 2006-07 -2013-2014 (or latest)

KS5 01-02 to 13-14, Tier 4, AY: 2006-07 -2013-2014 (or latest)

SC Schools 06-07 to 13-14 SUM, Tier 3, AY: 2006-07 -2013-2014 (or latest)

SLD-Level 3 VA, Tier 3, AY: 2006-07 -2013-2014 (or latest)

SLD-School Census, Tier 3, AY: 2006-07 (or earliest year)-2013-2014 (or latest)

SLD-SCAP, Tier 3, AY: 2006-07 (or earliest year)-2013-2014 (or latest)

SLD-Geo NSPD-NSPL, Tier 3, AY: 2006-07 (or earliest year)-2013-2014 (or latest)

SLD-CFR, Tier 3, AY: 2006-07 (or earliest year)-2013-2014 (or latest)

SLD-Institution History, Tier 3, AY: 2006-07 (or earliest year)-2013-2014 (or latest)

SLD-Institution Link, Tier 3, AY: 2006-07 (or earliest year)-2013-2014 (or latest)

6	Intended use of data
----------	-----------------------------

a	What are the aims of your project or research?
----------	---

The Academies Act, approved by the United Kingdom Parliament and implemented in July 2010, has led to a fast growth in the number of Academies. This has resulted in deep changes in England's public education provision. New Academy schools are opening and traditional public schools are converting into this new status, which gives them greater flexibility and decision-making abilities. A recent report by the Department for Education provides additional evidence that, in addition to an increase in numbers, academies also plan to expand in size (Cirin 2014).

The effects of these changes have been the object of a large public debate. Despite of this, to our knowledge the consequences of this reform on students have not been empirically documented (Machin and Vernoit, 2011, and Machin and Silva, 2013, study the effect of academy conversion on students focusing on the period 2001-2009). The primary purpose of our research project is to quantify the effect of the growth in the number and size of Academies following the 2010 Act on students' academic achievement and quality of educational experience.

In addition to studying the effects of this policy on children directly enrolled in Academies, we will also explore potential indirect effects that students attending traditional mandated schools might experience. We believe this margin can be important because the opening of an academy changes the education landscape that students (and parents) living in a given geographical area experience.

First, given that some students will switch from a mandated school to an Academy, the composition of the student body in the former will change. This could have consequences on the students who remain in mandated schools, through changes in each student's peers.

Second, the opening of an Academy (or conversion of a mandated school) might lead nearby mandated schools to lose pupils. This may alter the practices of mandated schools, which faced very little competition towards attracting students before this reform. This change in behaviour can have important effects on students' educational experience.

Lastly, the movements from mandated schools to Academies and the compositional changes that we described for students might also affect the teaching body. This is especially relevant given that Academies enjoy greater flexibility than mandated schools in setting their employees' salaries. This change in composition would affect both students enrolled in Academies and those enrolled in nearby mandated schools, with potentially large effects on learning opportunities of these

two groups. We plan on using teacher-level data from the School Workforce Census to document these movements and study this last channel.

b **Is there a specific question you are seeking to answer?**

The main questions we seek to answer are the following:

1. What is the effect of the growth in the number and size of Academies following the 2010 Act on students' academic achievement and quality of educational experience?
2. Are there indirect effects of Academies opening on students attending traditional mandated schools?
3. Does the composition of the teaching body in mandated schools change after the opening of an academy in its proximity?
4. Are the effects of Academies opening heterogeneous across students from different backgrounds and with different characteristics (SEN, minority, age)?
5. What determines a school's decision to convert into an Academy, in terms of previous student behaviour and achievement, financial conditions, capacity and enrolment? What are the consequences of the conversion of a mandated school into an Academy on the same variables?

c **Who is the intended audience?**

The intended audience of the proposed research is academia, specifically researchers in the fields of economics and education. We are happy to share our results with interested UK policymakers, as well as conducting additional analyses to answer related questions that might be useful or interesting to the Department for Education. We are available to travel to London to meet with DfE officials who might be interested in this project.

d **Why do you need the data requested to complete the aims of your project or research? You must specify why it is necessary for you to have the data you have requested including if appropriate whether you are conducting research or analysis, producing statistics or providing information, advice or guidance and how this will work will promote the education or well-being of children in England.**

- To identify the effects of Academy openings on the outcomes students enrolled in Academies, we will focus on mandated schools that convert into Academies. We will compare students who were already enrolled in a school before it converted with similar students in similar non-converting mandated schools in the same area. We will conduct these comparisons by estimating linear regression models of student achievement over time (before and after an academy conversion) as a function of individual level covariates as well as the type of school they attend (converted or not). To perform this analysis we require student level data on achievement, location of residence, school attendance and demographics.
- To identify effects on mandated schools students, we compare similar students in a location where an academy opens, with those in a similar location where no Academy opens, before and after the opening. The empirical model that we will estimate is a linear regression models of student achievement over time (before and after an academy conversion) as a function of individual level covariates as well as the type of location where they attend school (with or without a nearby Academy). To perform this analysis we require student level data on achievement, location of residence, school attendance and demographics.
- To identify the effects on teacher composition we would like to use rich teacher level data to infer what type of teachers (in terms of education, experience) are more likely to be recruited by academies after they open. To do so, we will estimate a linear probability model of the probability of working in an Academy after it opens as a function of teacher characteristics such as educational attainment, experience, age and gender. In addition to this we want to understand whether the probability of a teacher working in an Academy as opposed to a mandated

school depends on different salary opportunities by including salary as an explanatory variable in the above model.

Please see answer e. for details on why each sensitive variable is necessary for our research project.

This research will be beneficial for children in England because it will provide empirical evidence on the effects of a recent and highly debated education reform. Specifically our results can be useful in informing future policy decisions as they will shed light on the following:

- Effectiveness of Academies versus mandated schools in promoting student success as well as differential effects for different types of students. Together, this will inform on the desirability (or lack thereof) of further expanding the Academies program and what types of students to target. For example, were we to find that students from low socioeconomic backgrounds are the ones who benefit the most from attending an Academy, a policy recommendation would be to prioritize the opening of Academies in low socioeconomic status locations.
- Sorting of teachers between Academies and mandated schools based on observable characteristics. For example, were we to find that the most able teachers are moving to Academies at the expense of mandated schools, worsening the prospects of children in the latter, a policy recommendation could be to provide mandated schools with the appropriate tools to recruit valuable individuals.
- Identification of the characteristics of successful academies. Our data intensive project could highlight patterns that successful Academies follow. These could be used as recommendations for future converting schools, which would benefit their future students.

e Where fields constituting sensitive or disclosive data are requested, please explain on a field by field basis why each item is required and why the same research outcome cannot be achieved by using less sensitive or disclosive data

PLASC_Census 01-02 to 13-14 SUM, Tier 2, 4, AY: 2001-02 (or earliest year)-2013-2014 (or latest)

- AgeAtStartOfAcademicYear: This variable will allow us to take age effects into account in our regressions.
- EthnicGroupMajor: This variable will allow us to explore the heterogeneous effects of competition on students from different ethnic background, with a specific focus on ethnic minorities.
- FSMeligible: This variable will allow us to explore the (potential) differential effect of competition on students from different socioeconomic backgrounds, with a particular focus on disadvantaged students.
- EntryDate, LeavingDate: These variables will allow us to study student movements across schools as a result of competition. In addition to this, it will inform us about the amount of the exposure of each student to each school and its characteristics over time.
- SENprovision: This variable will allow us to explore the (potential) differential effect of competition on students with learning difficulties and/or disabilities.
- DistCurrSch, DistNearSch, NearSchLAEstab: These variables will be key to the analysis of school competition, as distance is one of the main determinants of school choice.
- LSOA: The specific location of each child is fundamental for the study of competition effects on children. In particular, this will allow us to determine how choices about schools are made in an environment in which the number and type of schools available to each child at a given distance changes over time. To do this, we need to determine the set of available schools separately for each child and within a given radius of distance. For this reason, variables indicating the distance to closest school and attended school, and the specific school attended, are not enough to perform this type of analysis. In addition to this, it should be noted that information on each student's postcode does not directly allow us to determine the set of schools available to each child in a reasonable distance radius, as these geographic units are very large.

f Are you proposing to match the data you have requested with any other data? (Please insert x)

Yes - Please set out your intentions and provide copies of any approval(s) you have obtained:		
No We are applying to obtain access to the School Workforce Census in order to answer our research questions having to do with teachers. However, we do not intend and we do not require to match NPD and SWFC data.		x
g	If the data is to be published or reproduced (having applied agreed disclosure control measures as in section k below) what format will this be in? (Please insert an x for all that apply)	
Academic research paper		x
Internal publication or report		
Website general access		
Website with restricted access		
Commercial publication for which no charge will be made		
Commercial publication for which a charge will be made		
Other (please detail)		
h	Details of other individuals who you propose to have access to the data including their name, job title, organisation and address and why it is necessary for them to have access. Please note each user is required to make an individual declaration and may be required to complete a separate Information Security Questionnaire.	
	Name	Job Title
	██████████	Graduate Student
	██████████	Graduate Student
	██████████	IT Services Officer
i	Where the individuals listed in “h” are from a different organisation to the applicant is there a contract in writing between the applicant and the other organisation(s) and does this cover data protection and information security arrangements?	
Yes - Provide a summary of the contractual relationship and the data protection and information security arrangements in place.		

No - Explain why no written contract is required and provide us with a summary of the data protection and information security arrangements in place?		
Not applicable - No additional organisations are listed in “h”.		X
j	When will you commence using the data? How long are you seeking to retain the data for? Why is this period necessary?	
We will commence using the data as long as it is made available to us. We seek to retain the data for a period of one year, in order to have the time to analyse it, to produce a research article, and to undergo the rounds of revision necessary for a refereed publication. This timeline is standard for the publication in Economics academic journals.		
k	Are you proposing to use the Standard Disclosure Control as set out at section 4.3 of the NPD User Guide? (Please insert X)	
Yes		X
No - How will you ensure that the confidentiality of individuals is preserved in any outputs arising from your use of the data?		

7	Have you discussed your request in advance with the DfE? If so, please specify the members of staff who have assisted you.	
Yes, [REDACTED].		

References

Cirin, R. 2014. “Do Academies Make Use of Their Autonomy?”, Department for Education, *Research Report*.

Machin, S., and J. Veroit. 2011. “Changing School Autonomy: Academy Schools and Their Introduction to England’s Education”, CEE Discussion Paper. No. 123. London School of Economics.

Machin, S., and O. Silva. 2013. "School Structure, School Autonomy and the Tail". CEP Special Report.



Department
for Education

© Crown copyright 2013

You may re-use this document / publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v2.0. To view this licence, visit www.nationalarchives.gov.uk/doc/open-government-licence/version/2 or email: psi@nationalarchives.gsi.gov.uk.

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Any enquiries regarding this publication should be sent to us at www.education.gov.uk/contactus.

This document is also available from our website at [NPD Online](#).