

CODEBOOKS for replication files.

ARTICLE: "The Unequal Distribution of Opportunity: A National Audit Study of Bureaucratic Discrimination in Primary School Access"

AUTHORS: Asmus Leth Olsen, Jonas Høgh Kyhse-Andersen and Donald Moynihan

DESCRIPTION: variable codebooks for analysis data.

VARIABLES: data file: "df1.Rda" (35 variables)

id - unique identifier for units but not linking to original data files in order to be compliant with GDPR interpretation in a Danish context.

send_date - date the email was sent out.

resp_date - date a response to the email was received. "NA" indicates missing values because the data is not available. This is due to non-response on the initial email request.

days_to_resp - number of days from email sent to response received. "NA" indicates missing values because the data is not available. This is due to non-response on the initial email request.

t_mohammad - dichotomous, primary treatment indicator, 1 if email is randomly assigned to be sent from "Mohammad Osman", 0 if sent from "Peter Nielsen".

t_good - dichotomous, secondary treatment indicator, 1 if email is randomly assigned to indicate that the student is "diligent", 0 if email contains no specific reason for the transfer.

treatments - categorical, containing the same information as the two previous variables. 0 if the email is from "Mohammad Osman" and with a "diligent student" cue, 1 if email is from "Mohammad Osman" and with no student cue, 2 if the email is from "Peter Nielsen" and with a "diligent student" cue, 3 if email is from "Peter Nielsen" and with no student cue.

conf_receipt - dichotomous, 1 if response email is just a confirmation of receipt of the email, 0 otherwise. "NA" indicates missing values because the data is not available. This is due to non-response on the initial email request.

forwarding - dichotomous, 1 if response email is a notification on the email being forwarded, 0 otherwise.

resp - dichotomous, 1 if a response email was received, 0 otherwise.

resp_true - dichotomous, 1 if a response email was received, 0 otherwise. "conf_receipt" and "forwarding" are coded to 0.

waitinglist - dichotomous, 1 if response text contains an invitation to get a spot on the waiting list, 0 otherwise.

spot1 - categorical, 0 if email response provides a clear rejection of the school transfer request, 1 if email response provides an unclear response to the school transfer request, 2 if email response provides a clear accept of the school transfer request. Coded by "Coder 1" and used in primary analysis.

reject1 - dichotomous, 1 if email response provides a clear rejection of the school transfer request, 0 otherwise. Non-responses are also coded to 0. Coded by "Coder 1" and used in primary analysis.

unclear1 - dichotomous, 1 if email response provides an unclear response to the school transfer request, 0 otherwise. Coded by "Coder 1" and used in primary analysis.

accept1 - dichotomous, 1 if email response provides a clear accept of the school transfer request, 0 otherwise. Coded by "Coder 1" and used in primary analysis.

spot2 - categorical, 0 if email response provides a clear rejection of the school transfer request, 1 if email response provides an unclear response to the school transfer request, 2 if email response provides a clear accept of the school transfer request. Coded by "Coder 2" and used for reliability test.

reject2 - dichotomous, 1 if email response provides a clear rejection of the school transfer request, 0 otherwise. Non-responses are also coded to 0. Coded by "Coder 2" and used for reliability test.

unclear2 - dichotomous, 1 if email response provides an unclear response to the school transfer request, 0 otherwise. Coded by "Coder 2" and used for reliability test.

accept2 - dichotomous, 1 if email response provides a clear accept of the school transfer request, 0 otherwise. Coded by "Coder 2" and used for reliability test.

simple_q - dichotomous, 1 if email response contains one or more simple questions, 0 otherwise.

simple_q_alt - dichotomous, 1 if email response contains one or more simple questions or in case of a non-response, 0 otherwise.

complex_q - dichotomous, 1 if email response contains one or more complex questions, 0 otherwise.

contact_q - dichotomous, 1 if email response contains a request to contact by phone, 0 otherwise.

meeting_q - dichotomous, 1 if email response contains a request to meet in person, 0 otherwise.

no_greet - dichotomous, 1 if email response contains no greeting and no name on the initial email sender, 0 otherwise. Non-responses are coded to 1.

greet_no_name - dichotomous, 1 if email response contains a greeting but no name on the initial email sender, 0 otherwise.

informal_greet - dichotomous, 1 if email response contains an informal greeting, e.g. 'hi', and name of the initial email sender, 0 otherwise.

formal_greet - dichotomous, 1 if email response contains a formal greeting, e.g. 'dear', and name of the initial email sender, 0 otherwise.

school_type - dichotomous, 1 if the school is a public school, 0 otherwise.

bigcity - dichotomous, 1 if the school is placed in one of the four major urban areas (i.e., Copenhagen, Aarhus, Aalborg, or Odense), 0 otherwise.

region - categorical, indicate what region of Denmark the school is placed in: Region Hovedstaden, Region Midtjylland, Region Nordjylland, Region Sjælland or Region Syddanmark.

above_median_student - dichotomous, 1 if the school has a number of students above the median level for the sample, 0 otherwise. "NA" indicates missing values because the data is not available. Newly opened schools will not have available data on the number of students.

above_median_grades - dichotomous, 1 if the school has a grade average above the median level for the sample, 0 otherwise. "NA" indicates missing values because the data is not available. This

can be due to two reasons. First, newly opened schools will not have available data for student grades. Second, if less than five students have participated in the exams then data is not released by the Ministry of Children and Education in order to safeguard anonymity of the students.

above_median_non_danes - dichotomous, 1 if the school has a share of students with non-native Danish background above the median level for the sample, 0 otherwise. "NA" indicates missing values because the data is not available. This can be due to two reasons. First, newly opened schools will not have available data for student grades. Second, if less than five students have non-native Danish background then the data is not released by the Ministry of Children and Education to safeguard anonymity of the students.

VARIABLES: data file: "school_df.Rda" (3 variables)

num_student - numeric, number of students enrolled in the school, 0 otherwise. "NA" indicates missing values because the data is not available. Newly opened schools will not have available data on number of students.

avg_grades - numeric, average grades of students enrolled in the school, 0 otherwise. "NA" indicates missing values because the data is not available. This can be due to two reasons. First, newly opened schools will not have available data for student grades. Second, if less than five students have participated in the exams then the data is not released by the Ministry of Children and Education to safeguard anonymity of the students.

pct_non_danes - numeric, percentage of students that are foreign-born or with foreign-born parents, 0 otherwise. "NA" indicates missing values because the data is not available. This can be due to two reasons. First, newly opened schools will not have available data for student grades. Second, if less than five students have non-native Danish background then the data is not released by the Ministry of Children and Education to safeguard anonymity of the students.

VARIABLES: data file: "responder_df.Rda" (4 variables)

sender_male - dichotomous, 1 if the response in the mail is sent

from a person with a male name, 0 otherwise. "NA" indicates missing values because the data is not available. This is typically the case if no name was provided in the response email or if the responder's gender could not be determined by the name.

sender_pos - categorical, 1 if the school principal responds, 2 if some other manager responds (e.g., vice principal), 3 if a secretary or assistant responds, 6 if the sender's position cannot be determined from email response. "NA" indicates missing values because the data is not available.

sender_eth - categorical, 1 if the response in the mail is sent from a person with a Muslim name, 2 if the name is a native Danish name, 3 if the name does not fit the categories of 1 or 2, 4 if the response does not provide a name. "NA" indicates missing values because the data is not available.

above_median_student - dichotomous, 1 if the school has an above median level number of students, 0 otherwise. "NA" indicates missing values because the data is not available. This can be due to two reasons. First, newly opened schools will not have available data for student grades. Second, if less than five students have participated in the exams then the data is not released by the Ministry of Children and Education to safeguard anonymity of the students.

EXTERNAL DATA SOURCES:

School level data from the year 2015/2016 used in all three data files:

Citation. Ministry of Children and Education. 2020. Educational Statistics - Excel access to the primary school area.

https://uddannelsesstatistik.dk/Pages/grundskolen_excel.aspx
[Retrieved October 7, 2020]

Instructions. Steps for accessing the data (the webpage is in Danish only):

- For school level data on number of students and number of students with non-native Danish background: Click "elevtal" and then again on the link called "elevtal" which is equivalent to this link:
<https://uddannelsesstatistik.dk/Pages/Reports/1577.aspx>
[Retrieved October 7, 2020]. By clicking the plus signs,

- the number of students (elevtal) becomes visible for each primary school. Then choose "herkomst" in the "PivotTable Fields" to the right. Again, by clicking the plus signs the number of non-native Danish students becomes visible ("udenlandsk").
- For school level data on grade averages: Click on "Overblik" and then "grundskole" under "skoletal". Grades are then found in column D, "Karaktersnit_skoletal" which is equivalent to this link:
<https://uddannelsesstatistik.dk/Pages/Reports/1834.aspx>
[Retrieved October 7, 2020].