

Finger, Claudia; Solga, Heike; Elbers, Benjamin

Article — Accepted Manuscript (Postprint)

Social inequality in admission chances for prestigious higher education programs in Germany: do application patterns matter?

European Sociological Review European Sociological Review

Provided in Cooperation with:

WZB Berlin Social Science Center

Suggested Citation: Finger, Claudia; Solga, Heike; Elbers, Benjamin (2024) : Social inequality in admission chances for prestigious higher education programs in Germany: do application patterns matter?, European Sociological Review European Sociological Review, ISSN 1468-2672, Oxford University Press, Oxford, Vol. 40, Iss. 6, pp. 1013-1029, <https://doi.org/10.1093/esr/jcae024>

This Version is available at:

<https://hdl.handle.net/10419/323384>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Online Supplement

for article:

Social inequality in admission chances for prestigious higher education programs in Germany: Do application patterns matter?

Section A: Information on clustering method and solution

Section B: Additional tables

- Table B1: Comparison of (weighting) variables between full-population register data and survey data
- Table B2: Mean (SD) of clustering variables and applicants' qualification profile, overall and by application cluster
- Table B3: Performance, waiting time, aspirations, constraints, strategies, and support, overall and by application cluster
- Table B4: SES differences in aspirations, constraints, strategies, support and information
- Table B5: Application patterns by parental occupation
- Table B6: Logistic regressions on admission chances (AME), only applicants with no or short waiting times (< 10 semesters)

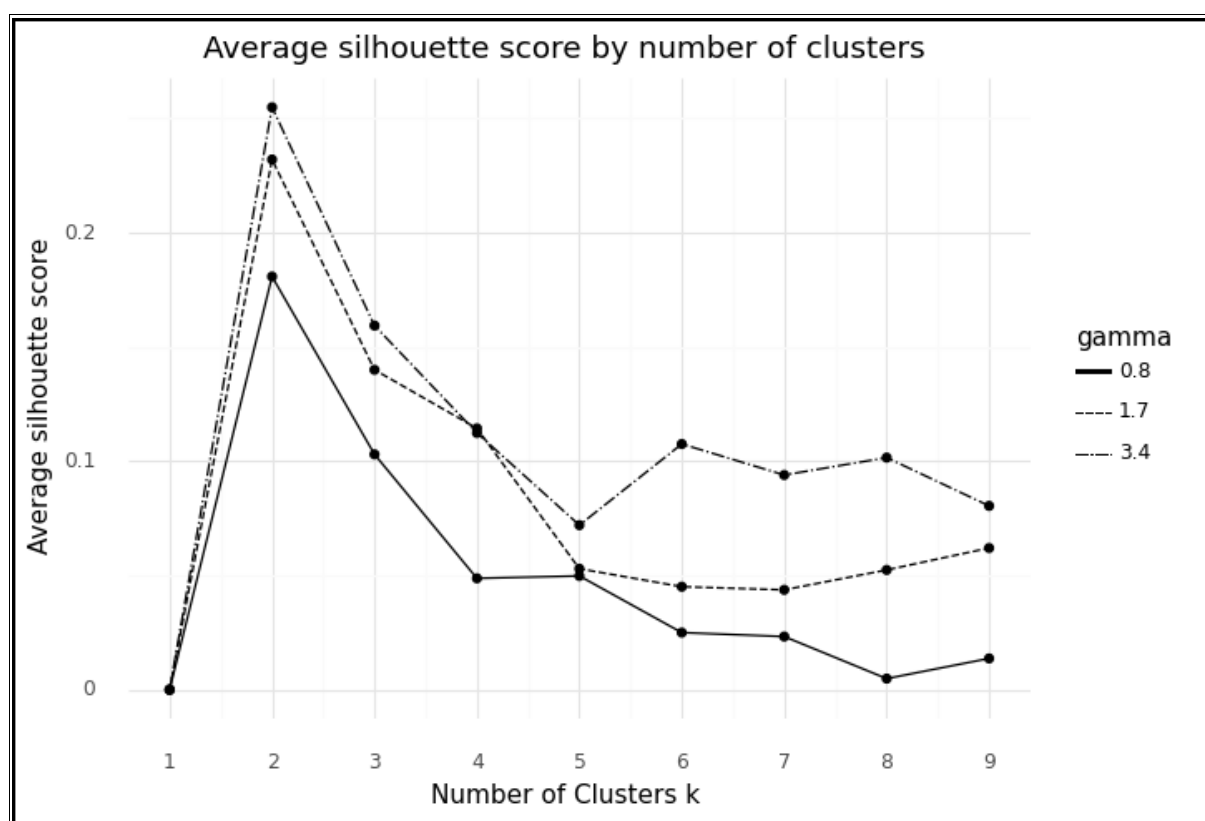
Section A: Information on clustering method and solution

The K-prototypes algorithm uses a combined objective function that includes distance metrics for both numeric and categorical variables. The parameter γ determines how far either type of variable is favored. When $\gamma=0$, then the clustering depends only on the numeric variables, which then results in the regular K-means algorithm.

We experimented with different values of γ to determine a good clustering solution. Figure A1 shows the average silhouette score by number of clusters and for varying values of γ . The silhouette score measures the distance to the other clusters for each observation. When the silhouette score approaches its maximum value of 1, it indicates that the observation is distanced from other clusters, indicating a well-separated observation. When the silhouette score is approaching its minimum of -1, then the observation is close to observations from other clusters, indicating that the cluster is not well separated. Good clustering solutions should have large average silhouette scores.

Figure A1 shows that, except for a spike around two to three clusters, the average silhouette score ranges roughly from 0 to 0.1 for different clustering solutions. These numbers indicate that we find a clustering solution with meaningful clusters, although these clusters still contain substantial heterogeneity and may overlap each other at the cluster boundaries. However, the silhouette score must be interpreted with caution for mixed data because its computation (as for many other evaluations of cluster fit) requires the definition of a distance metric. Silhouette scores use Euclidian distance as a metric, which is not appropriate for categorical variables. Evaluations of cluster fit must thus be interpreted with caution. We therefore did not base our choice of clustering solution entirely on the silhouette score but also qualitatively evaluated whether the clustering solution generated interpretable results. Ultimately, we settled on a clustering model with six clusters and $\gamma=3.4$.

Figure A1: Silhouette score by numbers of clusters



Source: SfH register data 2012-2018, authors' calculations.

Another possible way to evaluate the clustering fit is to assess whether cluster membership can be predicted using the clustering variables. This turns the problem into a “supervised learning” problem where the outcome values are known. To assess how successful such a model is, we split the data into a training and test data set and trained a gradient-boosted tree on the training data. We then evaluated how well the model predicted cluster membership on the test set (which was not used in the training step). The result of this analysis indicates that this model is very successful at predicting cluster membership on unseen data, as the F1-score is 0.99, just short of its maximum at 1, indicating perfect sensitivity and perfect positive predictive value.

Section B: Additional tables

Table B1: Comparison of (weighting) variables between full-population register data and survey data

	Application register 2018	Survey with weights (study sample)	Survey without weights (study sample)
<i>Observations</i>	42,285	3,470	3,470
Variables included in weights			
Gender (male)	32.9 %	32.9 %	27.4 %
Mean age (SD)	20.9 (3.5)	20.9 (3.4)	21.0 (3.4)
Mean GPA (SD)	1.9 (0.6)	1.9 (0.6)	1.7 (0.6)
Mean waiting time (in semester, SD)	3.6 (5.2)	3.9 (5.7)	4.1 (5.8)
Realistic waiting time	16.0 %	16.21 %	18.0 %
Admission (all quota)	24.4 %	24.7 %	42.6 %
Admission GPA quota	3.0 %	3.0 %	5.2 %
Admission waiting quota	4.8 %	4.8 %	6.3 %
Admission university-admission quota	16.7 %	16.7 %	28.8 %
Variables not included in weights			
Parental education			
0 college-educated parent		30.4 %	28.9 %
1 college-educated parent		30.5 %	30.5 %
2 college-educated parents		39.1 %	40.6 %
Test participation		41.9 %	48.3 %
Mean test scores (SD)		2.92 (0.6)	2.99 (0.6)

Source: SfH register data 2018, Online Survey of 2018 applicants to medical schools (waves 1 and 2), for weights see Section 4.1 in the main article, authors' calculations.

Table B2: Mean (SD) of clustering variables and applicants' qualification profile, overall and by application cluster (register data)

Cluster variables	C1 Constrained high achievers	C2 Constrained returners	C3 Strategic returners/ waiters	C4 Strategic preference maximizers	C5 Open searchers with strong preferences	C6 Open searchers with weak preferences	Total mean
2012-2018 (n = 274,838)							
Cluster size							
Absolute	35,357	40,258	42,393	43,033	45,666	68,131	
Relative	12.9%	14.6%	15.4%	15.7%	16.6%	24.8%	
# applications GPA quota	1.29 (1.31)	1.45 (1.95)	0.43 (1.03)	0.77 (1.75)	5.88 (0.60)	4.22 (2.60)	2.59 (2.69)
# applications waiting quota	0.91 (1.22)	1.72 (2.24)	4.77 (2.11)	0.66 (1.71)	5.09 (2.06)	5.04 (2.04)	3.30 (2.76)
# applications university-admission quota	2.00 (1.40)	1.09 (1.74)	5.28 (1.37)	5.19 (1.53)	5.56 (1.20)	5.72 (0.93)	4.39 (2.24)
First rank similarity	0.95 (0.22)	0.97 (0.18)	0.82 (0.38)	0.98 (0.13)	0.95 (0.23)	0.02 (0.15)	0.71 (0.46)
Quota similarity	0.83 (0.27)	0.95 (0.15)	0.44 (0.24)	0.93 (0.19)	0.69 (0.25)	0.34 (0.21)	0.66 (0.33)
Share of ranked test-based programs	0.86 (0.29)	0.06 (0.14)	0.63 (0.28)	0.66 (0.28)	0.60 (0.29)	0.69 (0.26)	0.59 (0.35)
Average distance first ranks (in km)	48.5 (65.1)	101.5 (133.5)	95.6 (119.7)	264.4 (189.9)	137.8 (156.2)	235.3 (133.9)	158.5 (159.5)
Maximum distance (in km)	108.2 (111.6)	224.2 (189.5)	381.1 (151.1)	430.9 (142.0)	407.2 (150.97)	497.9 (125.1)	364.1 (194.9)
Average reputation of first-ranked program	29.5 (24.5)	24.3 (23.2)	17.8 (18.6)	20.1 (21.3)	17.3 (19.0)	17.9 (15.5)	20.6 (20.4)
2018 only (year of online survey, n = 39,466)							
Cluster size							
Absolute	5,840	5,436	5,617	5,581	6,492	10,500	
Relative	14.8%	13.8%	14.2%	14.1%	16.5%	26.6%	
# applications GPA quota	1.37 (1.37)	1.53 (2.00)	0.45 (1.08)	0.81 (1.79)	5.90 (0.53)	4.35 (2.53)	2.72 (2.70)
# applications waiting quota	0.98 (1.28)	1.83 (2.31)	4.81 (2.06)	0.61 (1.63)	5.12 (2.02)	5.03 (2.05)	3.35 (2.74)
# applications university-admission quota	2.04 (1.40)	0.51 (1.19)	5.22 (1.41)	5.14 (1.54)	5.52 (1.29)	5.71 (0.96)	4.27 (2.32)
First rank similarity	0.94 (0.23)	0.97 (0.16)	0.82 (0.39)	0.98 (0.14)	0.95 (0.23)	0.03 (0.16)	0.69 (0.46)
Quota similarity	0.83 (0.27)	0.97 (0.12)	0.45 (0.23)	0.93 (0.18)	0.68 (0.24)	0.35 (0.21)	0.66 (0.33)
Share of ranked test-based programs	0.90 (0.25)	0.05 (0.14)	0.78 (0.21)	0.80 (0.18)	0.76 (0.24)	0.81 (0.20)	0.70 (0.34)
Average distance first ranks (in km)	46.6 (61.9)	109.4 (142.3)	88.0 (113.4)	247.3 (187.4)	128.8 (151.6)	226.5 (131.7)	150.9 (154.5)
Maximum distance (in km)	106.8 (110.9)	237.3 (199.3)	370.2 (156.8)	428.2 (145.4)	402.9 (154.4)	493.7 (126.8)	359.4 (198.2)
Average reputation of first-ranked program	27.5 (25.2)	28.0 (25.1)	17.2 (19.3)	20.6 (22.2)	17.5 (19.9)	17.9 (16.2)	20.9 (21.4)

Note: Standard deviation in parentheses. # means number.

Source: SfH register data 2012-2018, authors' calculations.

Table B3: Performance, waiting time, aspirations, constraints, strategies, and support, overall and by application cluster (online survey 2018)

	C1	C2	C3	C4	C5	C6	Total means	N ^{a)}
Cluster size	594 (16.8%)	367 (10.4%)	416 (11.8%)	732 (20.7%)	514 (14.6%)	908 (25.7%)		3,531
Key admission criteria								
Mean GPA (1.0–4.0) ^{b)}	1.55 (0.56)	2.25 (0.72)	2.17 (0.69)	1.84 (0.53)	1.87 (0.55)	1.97 (0.53)	1.93 (0.63)	3,524
Top GPA-quintile (%)	47.1	14.8	16.2	16.8	17.3	9.3	19.0	
Bottom GPA-Quintile (%)	8.3	40.3	29.9	12.6	13.0	14.6	17.9	
Test participation (TMS) (%)	32.2	13.1	40.5	47.4	42.6	55.0	41.7	3,478
Mean test score (if participated), z-standardized	0.26 (1.05)	-0.50 (0.81)	-0.29 (0.96)	0.17 (1.02)	-0.36 (0.97)	-0.19 (0.90)	-0.10 (0.90)	1,589
Mean waiting time (in semester) ^{c)}	2.8 (5.9)	7.5 (9.1)	5.2 (5.3)	1.8 (2.9)	2.9 (4.1)	4.3 (5.4)	3.9 (5.8)	3,470
Realistic waiting time (%)	11.2	42.0	27.1	2.7	10.8	15.5	16.0	
Aspirations and constraints								
<i>Is medicine your desired field of study? Since when? (%)</i>								3,501
No, not desired field	9.1	13.1	1.6	3.8	2.2	4.2	5.2	
Yes, since <15	38.9	37.4	39.8	47.1	52.5	45.8	44.3	
Yes, since 15-19	42.9	26.8	44.9	42.9	39.5	42.7	40.9	
Yes, since >19	9.1	22.7	13.7	6.2	5.8	7.3	9.7	
<i>Goal engagement: I am prepared to do everything that is necessary to study medicine (% responses items 4-5)^{d)}</i>								
Yes	80.4	78.8	90.9	87.1	91.4	91.3	87.4	3,506
<i>Medicine belongs to the most selective fields of studies. Would you be willing to ... (% items 4-5)^{d)}</i>								
Study a less selective field?	26.3	32.9	22.5	26.2	20.4	19.6	24.0	3,513
Study medicine at a location that is not your first choice?	37.3	56.4	73.7	83.9	81.2	91.7	74.5	3,513
<i>Reasons for choosing a university (% items 4-5)^{d)}</i>								
Proximity to home	74.6	60.0	55.4	29.1	46.5	30.9	45.5	3,519
Reputation	49.3	39.9	37.6	38.2	35.1	32.3	37.9	3,507
Admission chances in preferred field	42.6	55.7	70.3	79.3	77.9	85.2	71.7	3,481
Self-reported strategies								
<i>While applying, I selected and ranked universities in a way to increase ... (%)</i>								3,495
General admission chances	14.3	30.1	51.9	71.3	57.3	83.7	57.3	
Admission chances for preferred location	78.6	58.8	43.8	25.2	39.0	13.0	37.8	
<i>Reasons behind ranking in university-admission quota (% items 4-5)^{d)}</i>								
Match own profile – admission criteria	51.1	36.7	73.8	77.7	70.8	87.1	73.6	3,172
Strategic ranking of test-based programs	30.4	7.6	40.2	52.9	40.7	51.2	44.0	3,160
Ranking according to local preferences	74.2	70.0	51.6	40.2	51.1	28.1	46.2	3,176
<i>Including also alternative applications (%)</i>								
Medical programs abroad	5.0	9.1	15.8	17.7	16.1	24.5	16.2	3,497
Medical programs at private institutions	0.7	2.7	5.8	3.6	2.0	9.0	4.5	3,479
Non-medical programs	21.1	19.7	13.4	20.3	16.7	15.9	17.8	3,512
Support and information								
<i>Receiving support with applications (%)</i>								
By family members	48.1	37.4	43.0	52.1	50.1	43.8	46.4	3,502
Other persons	40.0	38.3	42.3	36.9	42.7	38.5	39.5	
<i>Knowing people with experiences in admission tests (%)</i>								
Family members	11.3	8.9	11.2	15.4	14.8	12.9	12.8	3,512
Other persons	55.1	63.5	65.3	64.9	61.7	68.8	64.0	

Notes: Standard deviations for GPA and test scores in parentheses.

- a) Case numbers per variable differ slightly due to missing values, mainly due to item non-response and do not exceed 1%. If the number of missing values is higher, this is due to routing decisions, i.e., certain questions have not been posed to all respondents (TMS scores only if participated; Ranking in university-admission quota only if applied via this quota).
 - b) Applicants' GPA potentially ranges from 1.0 (top) to 4.0 (bottom). The GPA quintiles include grades between: 1.0-1.3 (top quintile) and 2.5-4.0 (bottom quintile).
 - c) Applicants' waiting time ranges from 0 to 76 semesters. We defined "realistic waiting time" as at least 10 semesters. The waiting time necessary to get pre-selected for the waiting quota was 14 semesters in 2018. However, it is not a predefined threshold, but depends on the demand in a given year.
 - d) Responses were measured on a 5-point likert scale (1: not at all – 5: fully) and dichotomized (1-3=0; 4-5=1)
- Source: Online Survey of 2018 applicants to medical schools (waves 1 and 2), weighted (see Section 4.1 in the main article), authors' calculations.

Table B4: SES differences in aspirations, constraints, strategies, support and information

	Parental education			Total mean	N
	No college- educated parent	One college- educated parent	Two college- educated parents		
Aspirations and constraints					
<i>Is medicine your desired field of study, since when? (%)</i>					3,447
No, not desired field	5.1	4.8	5.2	5.1	
Yes, since < age 15	44.7	42.0	45.6	44.2	
Yes, since age 15-19	35.8	43.9	42.4	40.9	
Yes, since > age 19	14.5	9.3	6.7	9.9	
<i>Goal engagement: I am prepared to do everything that is necessary to study medicine (%) responses items 4-5)</i>					
Yes	87.6	88.5	86.3	87.4	3,452
<i>Medicine belongs to the most selective fields of studies. Under these conditions, would you be willing to (%) yes)</i>					
... study a less selective field	24.2	23.8	23.9	23.9	3,459
... study medicine at a location that is not your first choice	71.8	72.2	78.0	74.3	3,459
<i>Reasons for choosing a university....? (%) yes)</i>					
Proximity to home	51.5	47.9	38.8	45.4	3,465
Reputation	34.8	37.9	39.8	37.7	3,453
Admission chances	72.9	70.9	71.3	71.7	3,428
Self-reported strategies					
<i>While applying, I selected and ranked universities in a way to increase ... (%) yes)</i>					
... general admission chances	56.8	56.0	58.4	57.2	
...admission chances for preferred location	38.0	38.9	37.1	37.9	
<i>Reasons behind ranking in university-admission quota (%) yes)</i>					
Match own profile – admission criteria	72.2	72.2	75.1	73.4	3,123
Strategic ranking of test-based programs	38.2	44.2	47.9	43.9	3,111
Ranking according to local preferences	46.9	47.7	44.6	46.2	3,127
<i>Including also alternative applications (%) yes)</i>					
Medical programs abroad	10.4	16.0	21.1	16.3	3,444
Medical programs at private institutions	2.1	5.2	6.1	4.6	3,425
Non-medical programs	19.7	16.9	16.9	17.8	3,458
Support and information					
<i>Receiving personal support with applications (%) yes)</i>					
Family members	28.8	46.5	59.9	46.3	3,450
Other persons	38.8	38.3	41.6	39.7	
<i>Knowing people with experiences in admission tests (%) yes)</i>					
Family members	3.3	11.4	21.1	12.7	3,458
Other persons	58.9	65.5	67.2	64.2	
<i>Score (0-10) on knowledge quiz</i>					
Mean (SD)	5.85 (2.15)	5.94 (2.21)	6.19 (2.06)	5.98 (2.15)	3,470
N (observations)	1,003	1,058	1,409		3,470

Notes: Standard deviations in parentheses. Case numbers per variable differ slightly due to missing values, mainly due to item non-response and do not exceed 1%. If the number of missing values is higher, this is due to routing decisions, i.e., certain questions have not been posed to all respondents (variables: ranking in university-admission quota only if applied via this quota).

Source: Online Survey of 2018 applicants to medical schools (wave 1 and 2), weighted (see Section 4.1, main article), authors' calculations.

Table B5: Application patterns by parental occupation (in %)

Application patterns	Parental occupation		Total
	No parent working as physician	At least one parent working as physician	
Constrained high achievers (C1)	15.0	9.9	13.9
Constrained returners (C2)	12.0	9.0	11.4
Strategic returners/waiters (C3)	12.9	14.1	13.1
Strategic preference maximizers (C4)	20.0	24.8	21.0
Open searchers with strong preferences (C5)	13.8	15.9	14.2
Open searchers with weak preferences (C6)	26.4	26.4	26.4
<i>N (observations)</i>	<i>2,483</i>	<i>591</i>	<i>3,069</i>

Source: Online Survey of 2018 applicants to medical schools (wave 1 and 2), weighted (see Section 4.1, main article), authors' calculations.

Table B6: Logistic regressions on admission chances (AME), only applicants with no or short waiting times (< 10 semesters)

	M1	M2	M3	M4
Parental education (ref.: No college-educated parent)				
One college-educated parent	0.044* (0.020)	0.043* (0.019)	0.009 (0.014)	0.009 (0.014)
Two college-educated parents	0.092*** (0.019)	0.096*** (0.018)	0.026* (0.013)	0.024 (0.013)
Application pattern				
(ref.: Constrained high achievers (C1))				
Constrained returners (C2)		-0.301*** (0.033)		-0.061** (0.020)
Strategic returners/waiters (C3)		-0.195*** (0.034)		0.051* (0.021)
Strategic preference maximizers (C4)		-0.188*** (0.029)		0.041** (0.015)
Open searchers with strong preferences (C5)		-0.258*** (0.030)		-0.003 (0.018)
Open searchers with weak preferences (C6)		-0.280*** (0.027)		0.024 (0.016)
GPA and admission tests				
GPA-quintiles (ref.: top: 1.0-1.3)				
2nd: 1.4-1.6			-0.587*** (0.022)	-0.601*** (0.022)
3rd: 1.7-2.0			-0.702*** (0.019)	-0.717*** (0.019)
4th: 2.1-2.5			-0.750*** (0.018)	-0.764*** (0.018)
Bottom: 2.6-4.0			-0.745*** (0.020)	-0.759*** (0.020)
Test participation^{a)}				
(ref.: Nonparticipants)				
Participated with below-average scores			0.026 (0.014)	0.017 (0.014)
Participated with above-average scores			0.238*** (0.016)	0.222*** (0.015)
No information on test participation/scores ^{b)}			0.100** (0.034)	0.094** (0.033)
Pseudo- R ²	0.01	0.05	0.44	0.45
Observations	2,844	2,844	2,844	2,844

Average marginal effects (AME), standard deviations in parentheses.

^{a)} Only test for medical studies (TMS) considered. Four medical programs apply different test formats, but we do not have information on these tests in the data. If we include proxy measures, the findings do not change.

^{b)} 23 missing values for test participation and 55 for test scores.

Source: Online Survey of 2018 applicants to medical schools (waves 1 and 2), weighted (see Section 4.1, main article), authors' calculations.