



1

2 **Supporting Information for**  
3 **The Role of Professional Networks and Institutional Prestige in**  
4 **Shaping the First Career Moves of Scholars**

5 **Alexandra Rottenkolber, Ola Ali, Gergely Mónus, Jiaxuan Li, Jisu Kim, Daniela Perrotta, and Aliakbar Akbaritabar**

6 **Aliakbar Akbaritabar**

7 **E-mail: [akbaritabar@demogr.mpg.de](mailto:akbaritabar@demogr.mpg.de)**

8 **This PDF file includes:**

9 Figs. S1 to S3

10 Tables S1 to S12

11 SI References

## 1. Descriptive view of the data

Table S1 summarizes key characteristics of the study sample comprising 171,746 mobile scholars identified using Scopus-based bibliometric data. On average, individuals had published approximately 10 papers (SD = 20.51) and collaborated with over 15 first-order coauthors and more than 700 first- and second-order coauthors before their first move. The mean academic age at migration was 5.9 years (SD = 3.86). Institutional prestige, measured via the Leiden Ranking\*, was broadly distributed, with average target and source institution rankings around 3,800. The sample is skewed towards male scholars (66.7%) and is largely concentrated in Physical Sciences (44.6%), followed by Life Sciences (24.4%) and Health Sciences (19.7%), which is similar to the overall composition of scholars across fields of science in Scopus data (1). Most mobility events were national (55.3%), although international moves were also common (44.7%). The sample also spans a wide range of academic cohorts from 1996 to 2020, allowing to capture temporal variation in mobility patterns. Strength measures indicate that approximately 47.2% of scholars moved along at least one individual-level tie, while 46.0% moved along an institutional tie. Notably, only 19.1% had strong first-order-only individual connections, and 22.7% moved along alternative institutional pathways.

Variable	Overall (N=171,746)
<b>Cumulative Number of Publications</b>	9.97 (20.51)
<b>Number of First Order Coauthors</b>	15.36 (44.05)
<b>Number of First and Second Order Coauthors</b>	725.90 (2417.14)
<b>Career Age (Years)</b>	5.90 (3.86)
<b>Target Institution Ranking</b>	3,815.43 (3442.37)
<b>Source Institution Ranking</b>	3,908.10 (3307.17)
<b>Individual connection strength, continuous</b>	0.03 (0.08)
<b>Institutional connection strength, continuous</b>	0.01 (0.03)
<b>Gender</b>	
Female	57,185 (33.3%)
Male	114,561 (66.7%)
<b>Academic Cohort</b>	
1996-2000	24,514 (14.3%)
2001-2005	43,065 (25.1%)
2006-2010	59,687 (34.8%)
2011-2015	41,897 (24.4%)
2016-2020	2,583 (1.5%)
<b>ASJC Classification (Top)</b>	
Health Sciences	33,844 (19.7%)
Life Sciences	41,836 (24.4%)
Physical Sciences	76,665 (44.6%)
Social Sciences	19,109 (11.1%)
Unknown	292 (0.2%)
<b>Move Type</b>	
International move	76,814 (44.7%)
National move	94,932 (55.3%)
<b>Strength Measures</b>	
$S_{F,j}^{ind_{ait}}(T)$ Binary = 1	32,861 (19.1%)
$S_{F,j}^{ind}(T)$ Binary = 1	81,088 (47.2%)
$S_{i,j}^{inst}(T)$ Binary = 1	79,080 (46.0%)
$S_{i,j}^{inst_{ait}}(T)$ Binary = 1	38,985 (22.7%)

Table S1. Descriptive Statistics of the Study Sample

## 2. Decision Tree Results

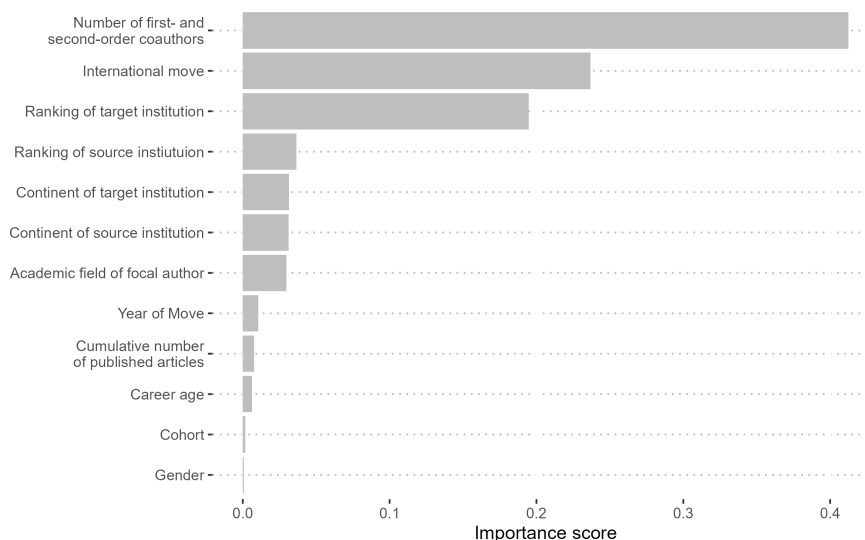
We constructed a decision tree to analyze the direction of scholarly mobility, classifying each move into one of four categories: along individual ties, along institutional ties, along both, or along neither. The model incorporates individual-level features (e.g., gender, career age, and connectedness) as well as contextual characteristics (e.g., move type and institutional rankings) to predict the direction of a scholar's movement. Our primary aim with this model was predictive—to identify the most important features that influence mobility direction.

\*<https://www.leidenranking.com/>

31 We trained the model using the `DecisionTreeClassifier` from Python’s `sklearn.tree` package (2). The decision  
 32 tree offers a transparent and interpretable framework, revealing both the hierarchy and relative importance of input  
 33 variables. Unlike more complex models, its structure allows for a direct understanding of how specific features  
 34 influence the classification outcome, i.e., the type of the move.

35 The dataset was randomly split into a 70% training set and a 30% test set. The classifier was configured with  
 36 `criterion= entropy`, `max_depth=10`, and `min_samples_leaf=10`, and class weights were balanced to mitigate class  
 37 imbalance (Table S2). The model achieved a weighted F1-score between 0.66 and 0.74 across categories (Table S4),  
 38 and an overall accuracy between 0.72 and 0.77 (Table S3), indicating a solid and balanced performance across the  
 39 four mobility pathways.

40 Among all input features, the size of a scholar’s second-order network—measured by the number of coauthors and  
 41 co-coauthors—emerged as the most influential predictor, contributing approximately 40% to the model’s decision-  
 42 making process (Figure S1). This finding underscores the critical role of extended personal networks in shaping  
 43 mobility patterns, suggesting that indirect connections (i.e., “friends-of-friends”) may be particularly associated with  
 44 the moves.



**Fig. S1. Decision tree.** The higher the importance score the more predictive a variable is for the type of mobility event observed (whether a move followed individual ties only, institutional ties only, both, or none of them).

45 The next most important features were the type of move (national vs. international) and the ranking of the  
 46 target institution. The importance of move type supports the idea that international mobility often follows different  
 47 mechanisms than domestic moves—potentially relying more on individual-level connections and less on institutional  
 48 pipelines. Likewise, the influence of institutional rankings suggests that prestige considerations weigh heavily in scholars’  
 49 decision-making processes, potentially guiding them toward highly ranked destinations when such opportunities are  
 50 available.

Parameter	Value
<code>class_weight</code>	balanced
<code>criterion</code>	entropy
<code>max_depth</code>	10
<code>max_features</code>	None
<code>min_samples_leaf</code>	10
<code>min_samples_split</code>	2

**Table S2. Decision Tree Parameters**

### 51 3. Alternative operationalization of the strength measures

52 Acknowledging that past collaborations only serve as a limited proxy for individual and institutional connections  
 53 (3, 4), we tested two additional measures to quantify individual and institutional connection strengths.

Direction	Accuracy
Following Individual Direction	0.72
Following Institutional Direction	0.74
Not Following Either Direction	0.74
Following Both Direction	0.77

**Table S3. Accuracy of the decision tree in prediction of each direction**

Direction	Precision	Recall	F1-score	Support
Following Individual Direction	0.70	0.75	0.72	24458
Following Institutional Direction	0.70	0.77	0.74	23760
Not Following Either Direction	0.64	0.67	0.65	18060
Following Both Direction	0.58	0.76	0.66	14723
Micro avg	0.66	0.74	0.70	81001
Macro avg	0.65	0.74	0.69	81001
Weighted avg	0.66	0.74	0.70	81001
Samples avg	0.65	0.69	0.63	81001

**Table S4. Classification Report of the decision tree for each direction**

54 **A. Alternative individual-level tie strength measure.** For the individual level metric, we also evaluated our analysis  
55 when taking only first-order ties into account. The corresponding measure  $S_{Fj}^{\text{ind}_{alt}}(T)$ , specified in Eq. 1, quantifies  
56 how strongly an individual  $F$  is connected to an institution  $j$  via their first-order co-authorship network in the year  
57 before a mobility event. It is defined as follows:

$$58 \quad S_{Fj}^{\text{ind}_{alt}}(T) = \frac{C_{Fj}^{1st}(T)}{\sum_j C_{Fj}^{1st}(T)} \quad [1]$$

59 where  $C_{Fj}^{1st}$  represents the number of first-order co-authors that individual  $F$  has at institution  $j$  at time  $T$  and  
60  $\sum_j C_{Fj}^{1st}$  is the total number of first-order co-authors that individual  $F$  has at the same time point  $T$ , the year before  
61 the mobility event, regardless of their institution.

62 **B. Alternative institutional-level tie strength measure.** The alternative measure we tested for the strength of the  
63 institutional connection is based on the past mobility of scholars between institutions. Specifically, two institutions  
64 are considered linked if individuals migrated from one to the other in a given year. The mobility-based strength  
65 measure,  $S_{ij}^{\text{inst}_{alt}}(T)$ , between institution  $i$  and institution  $j$  at a given year  $T$  takes the following form (Eq.2):

$$66 \quad S_{ij}^{\text{inst}_{alt}}(T) = \sum_{t=T-5}^T \frac{M_{ij}(t)}{\sum_{j \neq i} M_{ij}(t)} \quad [2]$$

67 where  $M_{ij}(t)$  is the total number of individuals migrating between the two institutions and  $\sum_j M_{ij}(t)$  is the total  
68 number of individuals who migrated from or to institution  $i$ , excluded those who migrated from or to institution  $j$ .

69 **C. Comparison of measures.** The two measures for the institutional level, where one is based on past collaborations  
70 and one on past mobility events, could be considered different proxies for the influence an institution has on the  
71 individual migration direction. For the main analysis, we took  $S_{ij}^{\text{inst}_{alt}}$  as it covers more percentage of the individuals  
72 (Table S5).

Strength measure	Individual-level connection strength		Institutional-level connection strength	
	$S_{Fj}^{\text{ind}_{alt}}$	$S_{Fj}^{\text{ind}}$	$S_{ij}^{\text{inst}}$	$S_{ij}^{\text{inst}_{alt}}$
% of individuals	22.0%	45.0%	35.0%	18.7%

**Table S5. Percentages of individuals who moved along a certain connection.**

73 **4. Multinomial Logistic Regression Models**

74 As a robustness measure, we fit our network direction-based multinomial logistic regression models with all possible  
 75 measurement combinations. There are four models all together, combining the two main measure and two alternative  
 76 measure. Our results show that while effect sizes vary between the different models, the significances and directions  
 77 are consistent across them.

**Table S6. Regression table - Multinomial logistic regressions - Individual direction measured with first-order coauthorship connections ( $S_{Fj}^{ind_{alt}}$ ), institutional direction measured with aggregated past coauthorships between institutions ( $S_{Fj}^{inst}$ )**

	<i>Dependent variable:</i>		
	both (1)	individual_direction (2)	institutional_direction (3)
Gender (male == TRUE)	-0.027 (0.020)	-0.049** (0.021)	-0.010 (0.014)
National move (TRUE)	2.357*** (0.023)	-0.146*** (0.022)	2.403*** (0.016)
Career Age	-0.008 (0.005)	-0.031*** (0.005)	-0.004 (0.004)
Network size (Medium)	0.793*** (0.028)	0.679*** (0.027)	0.051*** (0.016)
Network size (High)	1.607*** (0.030)	1.349*** (0.030)	0.070*** (0.020)
Productivity (Medium)	0.109*** (0.026)	0.150*** (0.027)	-0.022 (0.016)
Productivity (High)	0.508*** (0.029)	0.524*** (0.030)	-0.034* (0.021)
log(Ranking target institution)	0.972*** (0.011)	0.090*** (0.010)	0.962*** (0.008)
log(Ranking source institution)	0.255*** (0.011)	-0.046*** (0.010)	0.427*** (0.008)
Cohort fixed effects	Included	Included	Included
Period fixed effects	Included	Included	Included
Field fixed effects	Included	Included	Included
Geographic controls	Included	Included	Included
Constant	-10.089*** (0.155)	-1.696*** (0.151)	-9.938*** (0.114)
Akaike Inf. Crit.	322,959.6	322,959.6	322,959.6

*Notes:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01. Cohort fixed effects include indicators for cohorts 2001-2005, 2006-2010, 2011-2015, and 2016-2020. Period fixed effects include indicators for move periods 2011-2015 and 2016-2020. Field fixed effects include indicators for Life Sciences, Physical Sciences, Social Sciences, and other fields. Geographic controls include indicators for current and previous continents (Asia, Australia, Europe, Europe/Asia, North America, and South America).

**Table S7. Regression table - Multinomial logistic regressions - Individual direction measured with first-order coauthorship connections ( $S_{Fj}^{indalt}$ ), institutional direction measured with aggregated past mobility of scholars between institutions ( $S_{Fj}^{instalt}$ )**

	<i>Dependent variable:</i>		
	both (1)	individual_direction (2)	institutional_direction (3)
Gender (male == TRUE)	0.014 (0.025)	-0.035** (0.016)	0.032** (0.016)
National move (TRUE)	2.742*** (0.036)	-0.065*** (0.017)	2.687*** (0.022)
Career age	-0.007 (0.007)	-0.025*** (0.004)	-0.014*** (0.004)
Network size (Medium)	0.735*** (0.036)	0.726*** (0.023)	-0.019 (0.019)
Network size (High)	1.528*** (0.039)	1.457*** (0.024)	-0.025 (0.024)
Productivity (Medium)	0.150*** (0.034)	0.129*** (0.022)	-0.020 (0.019)
Productivity (High)	0.412*** (0.038)	0.541*** (0.024)	-0.129*** (0.025)
log(Ranking target institution)	0.935*** (0.014)	0.088*** (0.008)	0.907*** (0.009)
log(Ranking source institution)	0.376*** (0.014)	-0.083*** (0.008)	0.573*** (0.009)
Cohort fixed effects	Included	Included	Included
Period fixed effects	Included	Included	Included
Field fixed effects	Included	Included	Included
Geographic controls	Included	Included	Included
Constant	-12.813*** (0.206)	-1.408*** (0.122)	-12.449*** (0.138)
Akaike Inf. Crit.	285,709.6	285,709.6	285,709.6

*Notes:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01. Cohort fixed effects include indicators for cohorts 2001-2005, 2006-2010, 2011-2015, and 2016-2020. Period fixed effects include indicators for move periods 2011-2015 and 2016-2020. Field fixed effects include indicators for Life Sciences, Physical Sciences, Social Sciences, and other fields. Geographic controls include indicators for current and previous continents (Asia, Australia, Europe, Europe/Asia, North America, and South America).

**Table S8. Regression table - Multinomial logistic regressions - Individual direction measured with first- and second-order coauthorship connections ( $S_{Fj}^{ind}$ ), institutional direction measured with aggregated past coauthorships between institutions ( $S_{Fj}^{inst}$ )**

	<i>Dependent variable:</i>		
	both (1)	individual_direction (2)	institutional_direction (3)
Gender (male == TRUE)	-0.050*** (0.016)	-0.047*** (0.016)	0.010 (0.017)
National move (TRUE)	2.795*** (0.019)	0.357*** (0.017)	2.238*** (0.019)
Career age	-0.020*** (0.004)	-0.022*** (0.004)	-0.009** (0.005)
Network size (Medium)	1.098*** (0.020)	0.948*** (0.020)	-0.174*** (0.019)
Network size (High)	2.057*** (0.023)	1.877*** (0.022)	-0.424*** (0.025)
Productivity (Medium)	0.182*** (0.019)	0.261*** (0.019)	-0.024 (0.019)
Productivity (High)	0.527*** (0.022)	0.622*** (0.021)	-0.059** (0.024)
log(Ranking target institution)	1.182*** (0.009)	0.285*** (0.008)	0.885*** (0.009)
log(Ranking source institution)	0.267*** (0.009)	-0.098*** (0.008)	0.452*** (0.009)
Cohort fixed effects	Included	Included	Included
Period fixed effects	Included	Included	Included
Field fixed effects	Included	Included	Included
Geographic controls	Included	Included	Included
Constant	-10.968*** (0.135)	-2.317*** (0.125)	-9.748*** (0.133)
Akaike Inf. Crit.	356,441.9	356,441.9	356,441.9

*Notes:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01. Cohort fixed effects include indicators for cohorts 2001-2005, 2006-2010, 2011-2015, and 2016-2020. Period fixed effects include indicators for move periods 2011-2015 and 2016-2020. Field fixed effects include indicators for Life Sciences, Physical Sciences, Social Sciences, and other fields. Geographic controls include indicators for current and previous continents (Asia, Australia, Europe, Europe/Asia, North America, and South America).

**Table S9. Regression table - Multinomial logistic regressions - Individual direction measured with first- and second-order coauthorship connections ( $S_{Fj}^{ind}$ ), institutional direction measured with aggregated past mobility of scholars between institutions ( $S_{Fj}^{instalt}$ )**

	<i>Dependent variable:</i>		
	both (1)	individual_direction (2)	institutional_direction (3)
Gender (male == TRUE)	-0.015 (0.019)	-0.056*** (0.014)	0.046** (0.022)
National move (TRUE)	3.254*** (0.026)	0.552*** (0.014)	2.490*** (0.029)
Career age	-0.020*** (0.005)	-0.022*** (0.004)	-0.022*** (0.006)
Network size (Medium)	1.115*** (0.024)	1.027*** (0.016)	-0.236*** (0.024)
Network size (High)	2.062*** (0.027)	2.042*** (0.019)	-0.510*** (0.034)
Productivity (Medium)	0.197*** (0.022)	0.247*** (0.016)	-0.024 (0.024)
Productivity (High)	0.432*** (0.025)	0.625*** (0.018)	-0.063** (0.032)
log(Ranking target institution)	1.196*** (0.011)	0.342*** (0.007)	0.833*** (0.013)
log(Ranking source institution)	0.420*** (0.011)	-0.110*** (0.007)	0.566*** (0.013)
Cohort fixed effects	Included	Included	Included
Period fixed effects	Included	Included	Included
Field fixed effects	Included	Included	Included
Geographic controls	Included	Included	Included
Constant	-13.816*** (0.163)	-2.416*** (0.107)	-11.514*** (0.171)
Akaike Inf. Crit.	320,691.3	320,691.3	320,691.3

*Notes:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01. Cohort fixed effects include indicators for cohorts 2001-2005, 2006-2010, 2011-2015, and 2016-2020. Period fixed effects include indicators for move periods 2011-2015 and 2016-2020. Field fixed effects include indicators for Life Sciences, Physical Sciences, Social Sciences, and other fields. Geographic controls include indicators for current and previous continents (Asia, Australia, Europe, Europe/Asia, North America, and South America).

78 **5. Discrete Choice Models**

79 Regression table for the discrete choice model. We estimated three different models: (1) on the full data, (2) on male  
 80 scholars only, and (3) on female scholars only. The latter two were performed to test for gender differences, but none  
 81 were observed.

**Table S10. Regression table**

	<i>Dependent variable:</i>		
	choice (1) all	choice (2) men only	choice (3) women only
$S_{Pj}^{ind}$ binary (TRUE)	-6.994*** (0.398)	-6.191*** (0.487)	-8.609*** (0.695)
$S_{ij}^{instDCM}$ binary (TRUE)	-1.128*** (0.247)	-1.223*** (0.298)	-0.911** (0.440)
$S_{Pj}^{ind}$ continuous	0.728*** (0.027)	0.669*** (0.033)	0.846*** (0.047)
$S_{ij}^{instDCM}$ continuous	0.480*** (0.018)	0.477*** (0.022)	0.488*** (0.033)
International move (TRUE)	0.569*** (0.053)	0.581*** (0.064)	0.524*** (0.097)
log(Ranking target institution)	-0.385*** (0.052)	-0.421*** (0.063)	-0.327*** (0.091)
Region (Asia)	-0.105*** (0.025)	-0.095*** (0.031)	-0.143*** (0.046)
Region (Australia)	0.170*** (0.028)	0.109*** (0.034)	0.288*** (0.049)
Region (Europe)	-0.245*** (0.025)	-0.303*** (0.031)	-0.129*** (0.045)
Region (North America)	0.208*** (0.025)	0.160*** (0.031)	0.304*** (0.046)
Region (South America)	-0.003 (0.030)	-0.106*** (0.037)	0.172*** (0.052)
$S_{Pj}^{ind}$ binary (TRUE):log(Ranking target institution)	0.101*** (0.038)	0.112** (0.046)	0.086 (0.066)
$S_{ij}^{instDCM}$ binary (TRUE):log(Ranking target institution)	-0.114*** (0.020)	-0.093*** (0.024)	-0.151*** (0.036)
$S_{Pj}^{ind}$ continuous:log(Ranking target institution)	-0.027*** (0.003)	-0.028*** (0.003)	-0.024*** (0.005)
$S_{ij}^{instDCM}$ continuous:log(Ranking target institution)	-0.025*** (0.002)	-0.025*** (0.002)	-0.026*** (0.003)
$S_{Pj}^{ind}$ binary (TRUE):log(Ranking source institution)	-0.034 (0.042)	-0.145*** (0.052)	0.186** (0.073)
$S_{ij}^{instDCM}$ binary (TRUE):log(Ranking source institution)	0.058** (0.024)	0.062** (0.029)	0.042 (0.043)
$S_{Pj}^{ind}$ continuous:log(Ranking source institution)	0.011*** (0.003)	0.020*** (0.004)	-0.006 (0.005)
$S_{ij}^{instDCM}$ continuous:log(Ranking source institution)	0.006*** (0.002)	0.005** (0.002)	0.008*** (0.003)
Observations	8,582,100	5,724,300	2,857,800
Log Likelihood	-482,282.900	-323,538.200	-158,550.700
LR Test (df = 19)	615,618.600***	401,570.700***	214,435.800***

Note:

\* p<0.1; \*\* p<0.05; \*\*\* p<0.01

## 6. Multicollinearity Diagnostics

To assess potential multicollinearity among the key predictors, we examined pairwise associations using a Spearman correlation matrix (Table S11). As expected in scientific career data, productivity and network size exhibit a moderate positive correlation ( $\rho = 0.51$ ), reflecting the well-documented linkage between publication output and collaborations. This association is also mechanically consistent with our operationalization of individual connection strength in Equation (1): higher productivity increases the likelihood of co-authorship events, which in turn expands an author's collaboration network, either directly through a larger number of first-order ties or indirectly through second-order ties.

Correlations between prestige measures (log current and previous rankings) and the other predictors are positive but modest in magnitude, and the correlation between current and previous rankings remains well below levels typically associated with problematic multicollinearity.

	Career Age	Productivity	Network Size	log (Ranking of target institution)	log (Ranking of source institution)
Career Age	1	0.390***	0.230***	-0.028***	0.022***
Productivity	0.390***	1	0.514***	0.029***	0.052***
Network Size	0.230***	0.514***	1	0.149***	0.159***
log (Ranking of target institution)	-0.028***	0.029***	0.149***	1	0.216***
log (Ranking of source institution)	0.022***	0.052***	0.159***	0.216***	1

**Table S11. Pairwise Spearman correlations between career age, productivity, collaboration network size, and log-transformed rankings of target and source institution.**

\*\*\*  $p < 0.01$

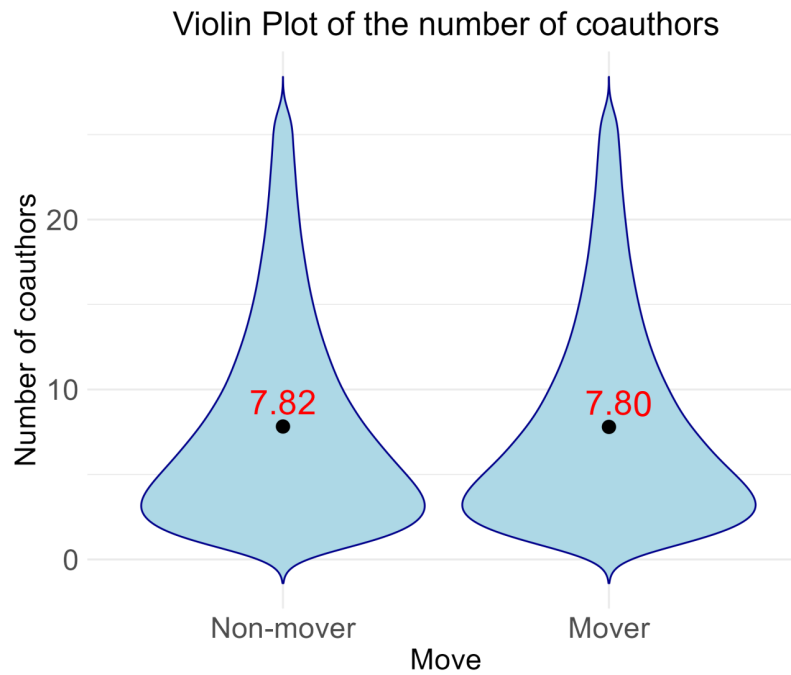
## 7. Matching Analysis: Co-authorship Network Size of Movers and Non-movers

To further assess whether the absence of a first mobility event is associated with differences in professional network size, we conducted an additional matching analysis comparing movers and non-movers with all observable characteristics available. Specifically, non-movers are defined as authors for whom no change in institutional affiliation was observed during the corresponding observation window following their first publication. To reduce imbalance in observable characteristics, we implemented a 1:1 nearest-neighbour matching procedure without replacement. Each mover was matched to one non-mover based on pre-move characteristics including gender, career age, academic field (Scopus ASJC classification), number of publications, region of the source institution, and academic cohort.

The outcome variable of this analysis is the size of the individual professional network, measured as the number of unique first- and second-order co-authors accumulated prior to the mobility event (for movers) or at the matched career stage (for non-movers). As shown in Figure S2, movers and non-movers exhibit nearly identical distributions of network size, with 7.82 co-authors for non-movers and 7.80 for movers on average, and the difference is statistically insignificant. These results indicate that, conditional on observable characteristics, non-mobility cannot be attributed to systematically smaller professional networks. The findings mitigate concerns that our mobility-conditional analyses disproportionately capture highly network-endowed scholars. However, it is also important to note that this analysis does not rule out the role of unobserved characteristics (e.g., talent, preferences, family constraints), nor does it imply that networks do not causally affect the probability of moving.

## 8. Relative Prestige Analysis

Prior work on faculty hiring networks has shown that upward moves are far less common than downward or lateral moves, reflecting a strict prestige hierarchy (5). We test whether analogous hierarchical patterns are present in early-career scientific mobility by constructing a relative prestige measure defined as the difference between the log prestige of the target and source institution. The Leiden Ranking provides multiple sorting criteria, including collaboration, open access, and scientific impact. Within the scientific impact dimension, institutions can be sorted by different indicators. We use P(top 50%) as it captures scientific impact without conflating prestige with the collaboration patterns we aim to explain or with open access considerations. The resulting distribution (Fig. S3) is approximately symmetric and centred around zero. To classify movements while avoiding arbitrary ordinal rankings, we adopt a conservative, distribution-based threshold approach: moves below  $-0.15$  standard deviations are classified as downward, moves above  $+0.15$  standard deviations as upward, and the remainder as lateral. The  $\pm 0.15$  standard deviations threshold corresponds to a small standardized difference in institutional prestige, ensuring that only moves that are meaningfully distinct from zero are classified as upward or downward.



**Fig. S2.** Comparison of co-authorship network size between matched movers and non-movers. The figure shows the distribution of co-authorship network size. Red numbers indicate mean values for each group.

123 Using this definition, approximately 41.7% of first career moves are upward, 12.4% are lateral, and 45.8% are  
 124 downward. These proportions differ markedly from faculty hiring networks, where upward moves are rare, reflecting  
 125 the fact that early-career mobility is not governed by the same hierarchical constraints that characterize senior faculty  
 126 hiring. Downward or lateral moves at this career stage often reflect labour market structure, postdoctoral placement  
 127 dynamics, and field-specific opportunity distributions rather than long-term stratification.

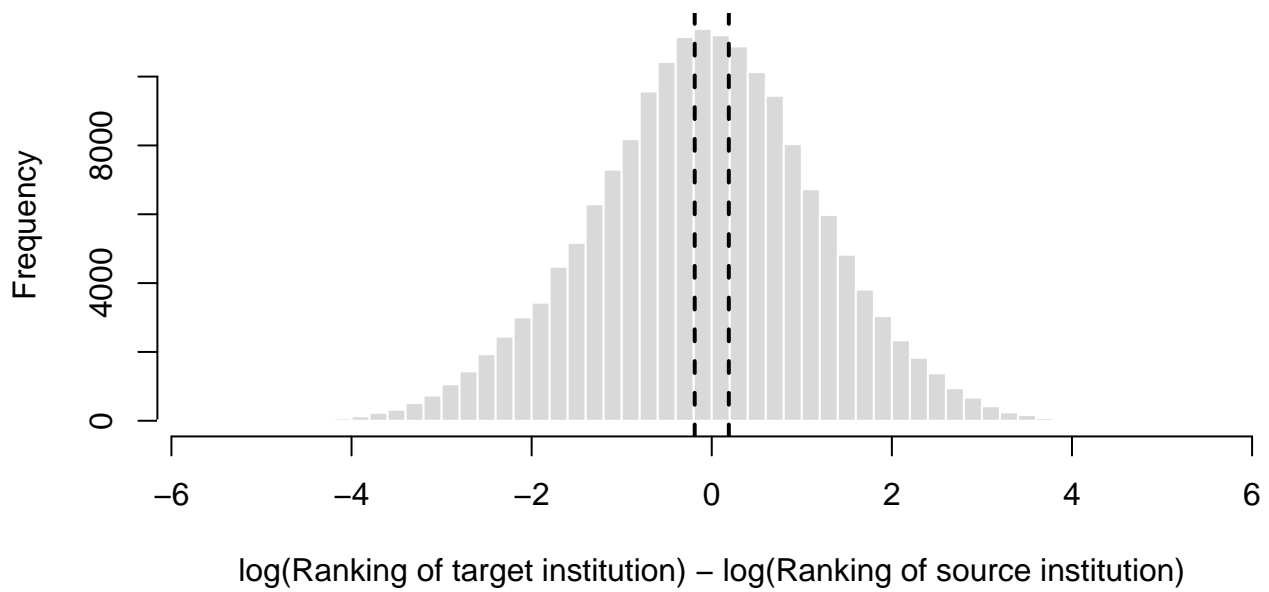
128 Table S12 reports the distribution of individual-level and institutional-level connections across relative prestige  
 129 classes. The prevalence of connection types varies only modestly across upward, lateral, and downward moves, and  
 130 we do not observe the strong hierarchical asymmetries reported in faculty hiring networks.

131 Together, these results indicate that relative prestige, as measured using the Leiden Ranking does not function  
 132 as the primary organizing principle for early-career scientific mobility in our data. We therefore retain absolute  
 133 institutional prestige as our main predictor, as it is better aligned with both the career stage under study and the  
 134 conceptual role of prestige in our analytical framework.

	Individual connection%	Institutional connection%	Both connection%	Neither connection%
Downward	18.9	16.5	23.4	41.2
Lateral	17.4	20.1	33.0	29.5
Upward	18.7	17.8	33.0	30.6

**Table S12.** Distribution of connection types exhibited by each relative prestige class.

### Distribution of Relative Prestige Change



**Fig. S3.** Distribution of relative prestige differences between destination and source institutions, measured as  $\log(\text{Ranking of target institution}) - \log(\text{Ranking of source institution})$ . The distribution is approximately symmetric and centred around zero. Vertical dashed lines indicate the  $\pm 0.15\sigma$  standard deviation thresholds used to classify moves as downward (below  $-0.15\sigma$ ), lateral (within  $\pm 0.15\sigma$ ), and upward (above  $+0.15\sigma$ ).

135

## References

136

1. J Baas, M Schotten, A Plume, G Côté, R Karimi, Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quant. Sci. Stud.* **1**, 377–386 (2020).

137

2. F Pedregosa, et al., Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

138

3. JS Katz, BR Martin, What is research collaboration? *Res. Policy* **26**, 1–18 (1997).

139

4. G Laudel, What do we measure by co-authorships? *Res. Eval.* **11**, 3–15 (2002).

140

5. A Clauset, S Arbesman, DB Larremore, Systematic inequality and hierarchy in faculty hiring networks. *Sci. Adv.*

141

**1** (2015).

142