Gender and Academic Mobility

This is a preliminary draft: please do not circulate without consent.

Aliénor Bisantis*

[Most recent version here]

November 10, 2025

ABSTRACT.

What explains the gender gap in academic careers? This paper studies the role of geographic mobility in shaping gender disparities. Using new administrative data from France linking the universe of PhD graduates to application and hiring outcomes in public universities, I track candidates from PhD graduation to permanent junior faculty positions. First, I show that women submit fewer applications, over shorter distances, and are more locally concentrated than men. Second, I exploit exogenous variation in the spatial distribution of job openings across disciplines and cohorts and estimate an instrumental variables model to recover the causal effect of mobility on hiring. Reduced mobility lowers both the number of applications and the probability of securing a position. Overall, these results suggest that geographic mobility constraints are a key mechanism behind persistent gender inequality in academia.

*Aix-Marseille University, CNRS, Aix-Marseille School of Economics; address: 5-9 Boulevard Maurice Bourdet, CS 50498 13205 Marseille Cedex 1; tel: +33 7 86 39 96 63; E-mail addresses: alienor.bisantis@univ-amu.fr

1 Introduction

Women now account for nearly half of all PhD graduates in Europe and the United States. Yet they remain persistently underrepresented among university faculty (She Figures 2021; NSF, 2023). Understanding why women "leak" from the academic pipeline is central not only for equity but also for efficiency. Academic careers are among the most selective and skill-intensive in the labor market, and losing female talent represents a large cost for both science diversity[1] and society. Why do these gender gaps persist?

A large literature has explored why women progress more slowly through academic careers. On the demand side, studies document hiring discrimination (Bagues et al., 2017a), recognition gaps in citation and peer evaluation (Sarsons (2017); Card et al. (2022)), and unequal access to professional networks (Ductor et al., 2023). On the supply side, women face hostile work environments (Wu, 2018), slower publication processes (Hengel, 2022), and strong family-related constraints (Kleven et al. (2019); Lassen and Ivandić (2024)). However, this work often abstracts from a defining feature of academic labor markets: the need for geographic mobility.

Mobility is often a prerequisite of academic careers. In many countries, tenure-track positions are scarce and geographically dispersed, forcing early-career researchers to apply broadly and relocate, sometimes multiple times. If women face higher mobility costs - due to family ties, dual-career constraints, or preferences for proximity Le Barbanchon et al. (2020) - then even equally productive women may apply to fewer institutions, over shorter distances, and face lower chances of obtaining permanent positions.

In this paper, I provide new evidence on how geographic mobility shapes gender differences in academic careers, using a unique combination of administrative, bibliometric, and geographic data for the universe of PhD graduates in France. I begin by documenting systematic gender differences in job search behavior: women apply to fewer institutions, over shorter distances, and are more likely to target universities near their PhD location. I then estimate the causal effect of mobility constraints on hiring outcomes, using an instrumental variable strategy that exploits variation in the spatial distribution of job openings across fields and years. The results show that reduced mobility significantly lowers the number of applications submitted and this subsequently lowers the probability of securing a permanent academic position.

The French institutional context offers a unique setting to study these questions. Recruitment into permanent university positions follows a centralized and highly transparent process. After completing a PhD, candidates must first obtain a national qualification to become eligible for a permanent junior faculty position (*Maître de Conférences*). Once

[1] Dossi (2024) shows that when smaller groups are underrepresented among researchers, this affects both the topics studied and the way research is conducted.

qualified, candidates apply simultaneously to openings posted by universities across the country - which makes geographic mobility a core part of the academic job search.

I construct a novel dataset that links three rich administrative sources: (i) the universe of doctoral theses defended in French universities (theses.fr); (ii) bibliometric data from Scopus, which I use to measure research productivity at the individual level; and (iii) application and qualification records from the Conseil National des Universités (CNU), which track eligibility, applications, and recruitment outcomes between 2009 and 2021. This dataset enables me to observe complete academic trajectories from PhD to hiring, link them to productivity, and characterize job search strategies in space.

I proceed in two steps. First, I document systematic gender differences in job search behavior. Women apply to fewer academic positions, over shorter distances, and to institutions located closer to their PhD university, consistent with stronger geographic constraints in their job search. These differences hold even when conditioning on field, cohort, and productivity. Second, I estimate the impact of these constraints on hiring outcomes using an instrumental variable design that exploits exogenous variation in the geographic distribution of job postings across disciplines and years. The results show that mobility frictions significantly reduce the likelihood of women obtaining a permanent academic position.

These findings highlight geographic mobility as a key, and previously underexplored, mechanism behind persistent gender inequality in academia. They suggest that reducing gender gaps in hiring requires addressing not only institutional biases or productivity differences, but also the structural features of academic labor markets. In particular, when job opportunities are scarce, rigidly timed, and geographically dispersed, they may disproportionately penalize candidates with higher mobility costs - even when those candidates are equally productive.

Related Literature This paper contributes to three strands of the literature on gender disparities in academic careers. A long-standing literature has documented women's persistent underrepresentation in academic careers, especially in STEM fields. Ginther and Kahn (2004) shows that women in economics face slower career progression, while Ceci (2011, 2014) and Meyer et al. (2015) emphasize both supply- and demand-side explanations. Huang (2020) provides cross-country evidence that gender disparities in scientific careers remain substantial despite near parity at entry. My contribution is to focus on the earliest stages of the post-PhD pipeline, showing where and how women's careers diverge from men's in a centralized and transparent academic system. Several studies highlight disparities at specific stages of the academic career. Bosquet et al. (2019) shows that women are less likely to be promoted within French economics departments.

Sarsons (2017) and Card et al. (2022) demonstrate gender gaps in recognition for co-authored work and peer evaluation, while Gaule and Piacentini (2018) and Lerchenmueller and Sorenson (2018) examine how advisors and early publication trajectories shape career outcomes. In France, Corsini et al. (2022) analyzes PhD students' productivity, and Patsali et al. (2024) studies research independence. Other work has pointed to structural frictions: Bagues et al. (2017a) documents hiring discrimination, Ductor et al. (2023) shows network disadvantages, Hengel (2022) highlights longer review times for female-authored papers, and Wu (2018) documents hostile work environments. My paper complements this literature by showing that gender gaps appear already at the transition into the first permanent job, with the final hiring stage accounting for most of the disadvantage. A growing literature emphasizes the role of family responsibilities in shaping careers. Kleven et al. (2019) show that childbirth generates large and persistent earnings penalties; in academia, Antecol et al. (2018) find that parental leave policies affect tenure outcomes, and Lassen and Ivandić (2024) and Galván and Tenenbaum (2024) document long-run penalties to mothers' careers. These family constraints are closely related to geographic mobility. Le Barbanchon et al. (2020) shows that women in the labor market often trade wages for shorter commutes. Few studies provide systematic evidence on mobility in academic job search. My paper is among the first to do so, showing that women apply to narrower and closer job sets, and that these mobility constraints help explain why women are less likely to secure academic permanent positions.

The remainder of this paper is organized as follows. Section 2 provides context on the French academic system. Section 3 describes the data sources and presents descriptive statistics. In Section 4, I document gender differences in application behavior and mobility. Section 5 documents the impact of application intensity on hiring outcomes. Section 6 concludes.

2 Institutional Context: The French Academic Pipeline

This section provides the institutional background to understand the structure of academic careers in France and how individuals progress from PhD completion to permanent positions. I first describe the organization of the French academic system, including the main ranks and recruitment procedures. I then present the structure of the academic pipeline, which outlines the key transitions from PhD graduation to permanent employment. The final part of the section summarizes three empirical facts that motivate the next stage of the analysis.

2.1 The French Academic System

This section describes how the French academic system works and presents the main stages from the PhD to a permanent position. The French system is highly structured, with national rules that apply to all universities. This organization makes it an ideal setting to study academic careers and gender differences in access to permanent positions. I first describe the two main faculty ranks that define academic careers, and then outline the steps leading from PhD graduation to a permanent junior position.

Faculty ranks and structure

University teachers and researchers in France are civil servants. There are two main ranks: *Maître de Conférences* (MCF), a junior permanent position, and *Professeur des Universités* (PR), the senior rank. The MCF is the first tenured position in a university, comparable to an assistant professor. The PR rank comes later through promotion and is similar to a full professor. Both combine teaching and research duties, with national rules for pay scales and promotion. The MCF rank therefore, represents the main entry point into a permanent academic career. this paper focuses on the transitions leading to that position. For clarity, I will refer to MCF positions as *junior permanent positions* throughout the paper.

From PhD to national qualification

After completing a PhD, candidates who wish to pursue an academic career must obtain a national *qualification* delivered by the National Council of the Universities (*Conseil National des Universités* - CNU)[2]. This step confirms that the person is eligible to apply for junior permanent positions. Applications are submitted online and include a CV, a list of publications, teaching experience, and other academic activities. Each discipline has its own CNU committee that reviews applications. The qualification is valid for four years, and candidates may apply in more than one disciplinary section (see Table A1 in the Appendix for an overview). Success rates range between 70 and 90% in most fields,[3] suggesting that this stage is not very selective.

[2]In rare cases, individuals from abroad may apply for university positions without the qualification, but this exception is uncommon. In practice, the qualification is almost always required to access junior permanent positions.

[3]The rate is around 50% in Law and Political Science due to a more restrictive selection policy.

From qualification to MCF recruitment

Once qualified, candidates can apply for junior permanent positions through the national online platform *Galaxie*[4]. All openings are published at the same time each spring, and candidates may apply to several universities. Each university establishes a selection committee composed of both internal and external members. Committees review applications, shortlist candidates, and conduct interviews. The process, therefore, combines national coordination with local autonomy. Junior permanent positions offer civil servant status, teaching obligations of 192 hours per year, research independence within a department, and job security. Promotion opportunities and salary progressions are uniform across universities, which limits within-rank inequality and facilitates comparisons across disciplines. 65% of qualified never apply to any position. In this paper, I will focus only on qualified candidates who have shown an interest in a position by looking at people who applied at least once.

Beyond the junior rank: Promotion and senior ranks

This paper focuses on the early stages of an academic career, up to the first permanent position. Later in the career, promotion to the senior rank, *Professeur des Universités* (PR), requires an additional qualification called the *Habilitation à Diriger des Recherches* (HDR). The HDR certifies that a researcher can supervise PhD students and lead research projects. Promotion to a senior position follows a process similar to the earlier qualification and recruitment stages, with some institutional changes introduced in 2018. Bosquet et al. (2019) studies gender differences in the transition from junior to senior positions within the French academic system in Economics.

Alternative research careers in France

Some researchers in France work in national research organizations such as the *Centre National de la Recherche Scientifique* (CNRS). These positions focus mainly on research and usually do not include teaching. They are fewer in number and very competitive, but they offer an alternative to university careers[5].

[4] Starting spring 2026, the platform will change, now called *Odyssée*.

[5] In Economics and related fields, some institutions have also introduced tenure-track *Assistant Professor* positions that lead to a tenured post equivalent to the senior permanent rank, *Professeur des Universités*.

2.2 French Academic Pipeline

This section describes the main stages of the French academic pipeline, from PhD completion to securing a permanent position (see Figure 1). The pipeline is structured around four key transition points that determine career progression and can be divided into two main stages.

Stage 1 covers the period from PhD graduation to obtaining the national qualification, which is required to apply for permanent junior positions. This stage includes two steps: (a) the decision to apply for the qualification and (b) the success of that application. *Stage 2* runs from qualification to securing a junior permanent position. It also includes two steps: (a) applying for a permanent position and (b) the outcome of the recruitment process.

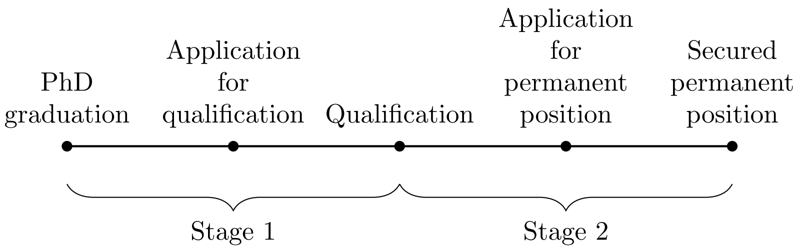


Figure 1: Key transition points

To analyze gender gaps at each transition stage of the academic pipeline, I rely on results developed in a companion paper Bisantis (2025). That work separates the overall probability of securing a permanent position into two components: (1) the probability of obtaining the national qualification, and (2) the probability of being hired into a junior permanent position once qualified. I briefly summarize the main findings here in Facts 1 and 2 to motivate the analysis of underlying mechanisms. The method is detailed in Section B.3 in Appendix.

Fact 1: There is a gender gap in access to permanent academic positions

Bisantis (2025) documents systematic gender differences in academic career progression. Across all fields, women are less likely than men to obtain a junior permanent position after completing their PhD. This disadvantage appears at multiple stages of the pipeline but is especially pronounced in the transition from qualification to recruitment into junior permanent positions.

Fact 2: The gap is driven by application behavior rather than success once applying

Using the decomposition approach, Bisantis (2025) quantifies how each transition contributes to the overall gender gap. The analysis shows that the main source of this gap lies in application behavior rather than in selection once candidates apply: Women and men have similar success rates conditional on applying, but women are less likely to submit applications for junior permanent positions. This finding motivates the next section, where I investigate one possible mechanism behind these application differences: geographic mobility.

3 Data

I built a novel dataset that links all PhD graduates in France to their research output, application behavior, and hiring outcomes in academia. This rich administrative data allows me to track candidates from PhD graduation to permanent junior faculty positions, observing both the universe of job offers and actual applications. I structure the data at the candidate–job level to analyze how mobility constraints shape application and hiring decisions. This section describes the data sources, variable definitions, and summary statistics. Summary statistics are reported in Table 4.

3.1 Sources

This study combines three main data sources to construct comprehensive academic career trajectories of PhD graduates in France: I retrieved data from *(i) Theses.fr* on all PhD theses defended in French universities from 2000 to 2021. For each PhD thesis, I have information on the discipline of study, defense year, university affiliation, and first and last name of the PhD graduate and the supervisor(s). *Theses.fr* is a centralized public platform with mandatory reporting from universities, making it a comprehensive and reliable source despite minor spelling inconsistencies and occasional delays. I infer supervisor gender using standard first-name-based classification methods, identifying gender for 95% of supervisors. I use *(ii) Scopus* to measure research productivity. I extract bibliometric data on all publications up to 2021, using full-name matching for both graduates and supervisors. This includes metadata on publication titles, journal names, year, number of co-authors, institutional affiliations, and the Article Influence Score (AIS)[6]. I use *(iii) CNU Database* from The Conseil National des Universités

[6]AIS is a journal-level metric commonly used to measure publication quality; see Bagues et al. (2017b).

dataset, which provides a comprehensive record of acceptance and rejection decisions for researchers seeking qualifications. It includes information on birth date, age, name, gender, and discipline associated with a candidate number. Using this candidate number, I track all applications to junior permanent positions and observe the selected candidate for each job.

3.2 Candidate–Job Dyad Construction

This study relies on a dyadic dataset that links each PhD graduate qualified to the universe of job openings for junior permanent positions in the French academic market in their discipline across institutions and years. The unit of observation is a dyad between a candidate and a potential job opening within the same academic field. To avoid selection bias, I restrict attention to the first year of job market participation.

Dyads are constructed by matching each PhD graduate to all job openings posted in their discipline during the relevant year. This approach reflects the actual opportunity set faced by candidates, as application rules and disciplinary boundaries limit cross-field mobility. Each dyad is associated with characteristics of the candidate, the job opening, and the hiring institution - including geographical distance between the PhD institution and the job-posting institution.

From this dyadic dataset, I construct an individual-level panel by aggregating across dyads. I exclude job postings in overseas territories (DOM-TOM) except Corsica and drop graduates or applications associated with those regions (representing 3% of the sample).

The final sample comprises approximately 2,287,593 dyads, constructed from 43,966 qualified PhD graduates and 18,787 job offers across 58 sub-disciplines, spanning the period from 2009 to 2021.

3.3 Main Variables

Outcomes For each dyad, I observe two main outcomes: *Apply* takes the value 1 if the candidate applied for the position, and 0 otherwise. This outcome is used to analyze revealed preferences over job openings. *Success* takes the value 1 if the candidate was selected for the position, and 0 otherwise. This variable captures hiring outcomes.

Controls The vector of controls is a function of age at the year of application and its square; whether individual i has at least one scientific publication appearing in the *Scopus platform* (dummy $Publish_{it}$); the cumulative number of publications at year t ($Quantity_{it}$) and the cumulative Article Influence Score (AIS) of publications at year t

($Quality_{it}$), and supervisor characteristics including whether at least one supervisor is female ($Female\_supervisor$) and the cumulative AIS score of supervisors at the year of PhD defense of individual i ($Quality\_supervisor_i$). All controls are listed in Table 4.

Distance The main geographic variable is the great-circle (orthodromic) distance between the city of the PhD-granting institution and the city of the hiring institution, measured in kilometers. I compute this distance "*à vol d'oiseau*" using geo-coordinates (latitude and longitude), following the Vincenty ellipsoid formula. This measure reflects true geographic separation, abstracting from travel infrastructure. To account for spatial variation in large urban areas and to improve precision in cases where both institutions are located in the same city, I follow the approach of Mayer and Zignago (2011) by incorporating the radius of the city. The city radius provides an estimate of the geographical size of each urban area and helps differentiate between genuinely proximate institutions and those that may be several kilometers apart within the same city.[7]

Distance as Proxy for Home Bias The use of distance from the PhD institution as a measure of spatial frictions is motivated by the concept of *home bias* - the well-documented tendency for individuals to remain near familiar or previously inhabited locations[8]. While I do not observe candidates' place of birth, the PhD institution serves as a reasonable proxy for "home" for several reasons: (1) Many candidates complete both their Master's and PhD at the same institution,[9]; (2) The PhD period often coincides with long-term personal and professional settlement; (3) Application patterns in the data show strong spatial concentration around the PhD institution - for instance, one quarter of applications are submitted within the same region. This interpretation is consistent with recent literature documenting geographic immobility and local labor market attachment: prior residence and institutional affiliation are shown to influence job search behavior (Kleven et al., 2020; Diamond, 2016).

[7] I use INSEE data to obtain the official surface area of each city and compute the radius assuming circular symmetry. Geo-coordinates (longitude and latitude) for each city are also retrieved from INSEE. I compute great-circle distances between cities using the *GEODIST* function in Stata based on these coordinates.

[8] The concept of *home bias* originates in international finance, where it describes the preference for domestic over foreign assets. It is now commonly used in labor and migration contexts to capture individuals' tendency to remain near familiar or previously inhabited locations.

[9] In France, Master's programs include research-oriented tracks that often serve as a direct pipeline to a PhD at the same university.

3.4 Descriptive Statistics

3.4.1 Candidates

Table 4 presents descriptive statistics for the sample of qualified candidates, disaggregated by gender. Women represent 44% of the sample, are slightly older than men on average (34 vs. 33 years), and have spent a similar amount of time since completing their PhD (3 years).

Male candidates are more likely to have at least one publication (64% vs. 49%), publish more (5 vs. 3), and have higher cumulative journal impact scores (AIS: 6 vs. 3). These differences are consistent with the literature on gender gaps in research productivity[10]. Both groups apply to a similar number of positions (3.5 applications), and the probability of securing a position is identical across genders at 9%.

Some of these differences likely reflect disciplinary composition rather than gender per se. Women are more represented in Humanities (31% vs. 18%) and Literature (12% vs. 5%), while men dominate Engineering (13% vs. 6%), Physical Sciences (15% vs. 7%), Mathematics (8% vs. 3%), and Computer Science (11% vs. 4%).

Field-level statistics (Tables B3–B6) show that gender gaps in research productivity and application behavior persist within-but vary across-fields. In STEM, men publish more and apply slightly more; in Humanities, women apply more but publish slightly less. In Biological and Social Sciences, gender differences are minimal. Age also varies considerably across fields: candidates in Humanities and Social Sciences are older on average than those in STEM or Biological Sciences, which likely contributes to the small overall gender difference in age, given women's greater representation in those older fields.

While the magnitude and direction of gender gaps are not uniform across fields, men tend to have higher research output on average, particularly in fields with greater overall publication intensity. These patterns indicate that differences in field composition alone are insufficient to explain all observed gender disparities, reinforcing the importance of including field fixed effects in the empirical analysis.

Successful Candidates. Table B2 presents descriptive statistics for candidates who secured a permanent position. Women account for 42% of this group. Among successful applicants, men continue to show higher research output, with more publications (4 vs. 2) and higher cumulative AIS scores (3.25 vs. 1.59). To secure the position, women applied to slightly more positions on average (12 vs. 11) and are, on average, marginally older (32 vs. 33 years). Gender differences in field representation persist: women remain

[10] @@

more concentrated in Humanities, Literature, and Management, while men are more represented in Engineering, Computer Science, and Physical Sciences.

Location Figure 3 displays the cumulative number of qualified candidates from 2009 to 2021 by *department*[11], based on the location of their PhD institution. The spatial distribution is highly unequal across France. A small number of departments concentrate the majority of qualified candidates. Paris alone accounts for over 14,500 qualifiers, representing nearly 30% of the national total. Other prominent academic hubs include Rhône (Lyon, 2,612), Haute-Garonne (Toulouse, 2,479), Isère (Grenoble, 1,897), Bouches-du-Rhône (Marseille, 1,995), and Hérault (Montpellier, 1,655). In contrast, more than 40 departments recorded fewer than 100 qualifiers, and over 30 produced none at all during the entire period.

Figure 4 shows the evolution of the number of qualified candidates per department between 2009 and 2021. The left y-axis plots the annual counts for each department (excluding Paris), while the right y-axis displays the national totals. Two versions of the total are shown: a solid line represents the sum excluding Paris, and a dashed line includes Paris. This distinction is necessary because Paris is a strong outlier and would otherwise obscure variation across other departments. The Figure highlights a clear national decline in the number of qualified candidates starting around 2014. Most departments follow a downward trajectory, though the decline is numerically driven by the largest academic centers-particularly Paris and other major university cities.

3.4.2 Job Offer

Location. Figure 5 shows the cumulative number of permanent academic job openings by *department* between 2009 and 2021, based on the location of the hiring institution. The spatial distribution broadly overlaps with the training locations of qualified candidates, but job openings are overall less concentrated. The department of Paris again dominates with over 4,500 positions, followed by Rhône (1,058), Haute-Garonne (892), Bouches-du-Rhône (749), and Isère (647). However, many other departments offer relatively few jobs: over 30 departments recorded fewer than 100 positions during the entire period, and more than 20 had none at all.

Figure 6 displays the annual number of permanent academic job offers by department between 2009 and 2021. Department-level trends are plotted on the left y-axis, excluding Paris for readability. The right y-axis shows two national totals: the solid line excludes Paris, while the dashed line includes it. As in the case of qualified candidates, the

[11] Departments (départements) are French administrative divisions, akin to counties, and serve as a geographic unit in the analysis.

number of job openings has declined significantly since 2014. However, the contraction in job supply is even more pronounced, with a steeper and more sustained decline. This reflects broader institutional constraints on recruitment and shrinking opportunities.

Figure B2 and Figure B1 document the evolution of both supply and demand in the French academic job market from 2009 to 2021. The number of available junior positions has declined steadily since 2012, across nearly all disciplines. This contraction has been met with relatively stable or increasing numbers of qualified candidates, suggesting a tightening of the market over time. Disciplines such as Humanities consistently offer the largest number of positions, but they also encompass a broader range of subfields (see Table A1 in the Appendix).

4 Gender gap in application behavior

4.1 Results: Dyad Approach to Application Behavior

To estimate how spatial frictions shape job search behavior, I examine the probability that a qualified PhD graduate applies to a given junior permanent position. The unit of observation is a dyad between candidate i and job opening j in discipline f during year t. The sample is restricted to the initial year of job market entry and to job openings within the candidate's discipline of qualification. I estimate the following linear probability model:

$$Y_{ijtf} = \beta_1 \ln(Distance_{ij}) + \beta_2 Female_i + \beta_3 \ln(Distance_{ij}) \times Female_i + X'_{ijtf}\gamma + \alpha_{itf} + \delta_{jtf} + \varepsilon_{ijtf} \quad (1)$$

The dependent variable is a binary indicator equal to one if candidate i applied to position j. The key independent variable is the log of the great-circle distance (in kilometers) between the PhD institution and the hiring institution. I interact $ln(distance)_{ij}$ with a gender dummy to test whether female candidates are more sensitive to spatial frictions. The vector X_{ijtf} includes controls for candidate age, publication record, and supervisor characteristics. Fixed effects α_{itf} and δ_{jtf} absorb candidate and job heterogeneity within discipline–year cells.

Table 1 presents the estimation results. Across all specifications, the interaction between distance and gender is negative and statistically significant, indicating that female candidates are more geographically constrained in their application behavior.

In Column (1), the specification includes university-by-year-by-field fixed effects for both the candidate's PhD institution ($U_i \times t \times f$) and the job institution ($U_j \times t \times f$). This specification compares candidates from the same university and discipline applying in the

Table 1: Determinants of Application Behavior: Candidate–Job Dyads

	(1)	(2)	(3)
Dependent variable:		*Apply to position*	
Female	0.00711**	0.000635	-
	(0.00280)	(0.00324)	-
ln(Distance)	-0.0127***	-	-0.0127***
	(0.000319)	-	(0.000304)
ln(Distance) $\times$ Female	-0.00238***	-0.00116**	-0.00235***
	(0.000440)	(0.000534)	(0.000398)
Adj R^2	0.19	0.19	0.30
Controls	yes	yes	yes
Fixed effects	$U_i \times t \times f + U_j \times t \times f$	$U_i \times U_j \times t \times f$	$i \times (t \times f) + j \times (t \times f)$
Observations	2,287,422	2,162,136	2,286,953

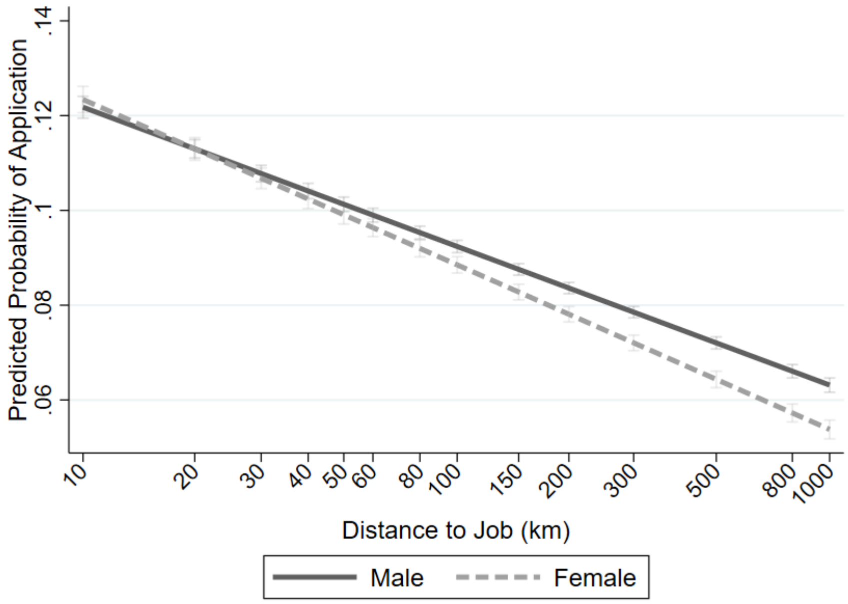
Notes: *Apply* is a binary variable equal to 1 if the candidate applied to a specific job. Each observation represents a dyad between a candidate and a potential job opening. *ln(Distance)* is the logarithm of the geographical distance between the job and the candidate's PhD institution. All regressions include controls for age, publication metrics, and supervisor gender. Fixed effects vary across specifications and are indicated in the "Fixed effects" row: U_i denotes the university of candidate i, U_j the university of the job j, t the year, and f the field. i and j denote candidate and job identifiers, respectively. Standard errors are clustered by Discipline $\times$ Candidate Univ $\times$ Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

same year, and jobs posted by the same hiring institution and discipline in the same year. The coefficient on $\ln(Distance_{ij})$ therefore captures how variation in distance across dyads - holding constant institutional characteristics - predicts application decisions. The coefficient on $\ln(Distance_{ij})$ is 0.0127, implying that a 10% increase in distance reduces the probability of application by about 0.13 percentage points. The interaction term is negative and significant (-0.00238), suggesting that this spatial sensitivity is amplified for women. The combined effect of distance for female candidates is 0.0151, nearly 20% larger in magnitude compared to men. Figure 2 illustrates this result: the probability of applying declines with distance for both genders, but the slope is notably steeper for women. In Appendix Figures and , I show non-parametric binscatters that flexibly depict the same pattern. These reveal that women are particularly less likely to apply beyond 200km, and the gap persists across specifications with richer fixed effects.

Column (2) introduces dyadic fixed effects at the candidate university–job university–year–field level $U_i \times U_j \times t \times f$. This specification compares candidates from the same PhD institution applying to jobs at the same hiring institution, within the same discipline and year. The main effect of distance is absorbed, but the interaction term remains negative and statistically significant (0.00116), confirming that gendered distance effects persist even within narrowly defined institutional pairs.

Column (3) includes the most restrictive specification, with individual-level fixed

Figure 2: Predicted Number of Applications by Distance to Job Offers by Gender



Notes: Predicted probability of applying to a job as a function of distance from the candidate's PhD institution, shown separately by gender. Estimates are based on the regression model in column (1) of Table 1 and control for age, publication metrics, supervisor characteristics, and fixed effects. Distance (x-axis) is plotted on the original kilometer scale for interpretability. Standard errors are clustered by Discipline X Candidate Univ X Year.

effects for both candidates and jobs, interacted with year and field $(i \times (t \times f) + j \times (t \times f))$. This approach compares which jobs a given candidate applies to, and which candidates apply to a given job, within the same field and year. By absorbing all individual and job-level characteristics - both observed and unobserved - that are constant within the year-field cell, this specification sharpens identification by leveraging only within-candidate variation in job opportunities. The remaining variation in distance captures differential application patterns across jobs faced by the same candidate. The distance coefficient remains negative and highly significant (0.0127), and the interaction term remains robust (0.00235). Even when comparing the same candidate across alternative jobs, and the same job across alternative candidates, women remain less likely to apply to geographically distant positions.

4.2 Robustness: Dyad Approach to Application Behavior

The results are robust to a range of alternative specifications. First, Table D12 replaces the logarithmic transformation of distance with the level measure (in kilometers). The interaction between gender and distance remains negative and significant across specifications, confirming that the log-linear form is not driving the result.

Second, I construct a new measure of geographic frictions based on estimated commuting time between the PhD and job location. This variable combines train travel time (from official SNCF timetables), road travel time (based on routing algorithms), and AI-based predictions for less connected pairs. Details of the construction are provided in Appendix Section D.2. As shown in Table D13, the interaction between gender and commuting time remains negative and significant.

Third, I include controls for age at PhD and time since graduation to account for potential differences in life-cycle stage (Table D18). The estimates are unchanged, suggesting that career timing is not a confounding factor.

Fourth, to assess whether the effect is driven by spatial clustering in the Paris region - where job opportunities are dense - Table D15 excludes candidates located in Paris. The gender-distance interaction remains robust, indicating that local agglomeration is not driving the main result.

Fifth, I account for potential selection based on the decision to apply at all. Table D16 presents estimates for several restricted subsamples. Panel A focuses on candidates who applied to at least one job during their entire career, while Panel B restricts further to those who submitted at least one application in their first year of eligibility.

4.3 Heterogeneity in Dyad-Level Results

To better understand the mechanisms underlying this gendered spatial constraint, I next examine how the distance penalty varies across key dimensions of candidate heterogeneity.

Table D18 explores heterogeneity in the gender-distance interaction by candidate age, time since PhD, academic productivity.

Age and Career Stage. Panels A and B split the sample by the median candidate age at application and years since PhD, respectively. The gender-distance interaction is negative and statistically significant in both younger and older groups, but the magnitude is larger among older candidates and those who are further from graduation. For example, among those with above-median time since PhD, the interaction term is -0.00288 compared to -0.00150 for newer graduates

Research Productivity Panels C and D examine heterogeneity by candidates' academic productivity, measured by AIS (Article Influence Score) and number of publications. The gender-distance gap is significant regardless of research output, but larger among those without any publications. For instance, women with no publications face a higher distance penalty (-0.00207) than their male peers, while those with publications still show a significant, but smaller, gap.

After First Year of Qualification. Panels C–E of Table D16 examine how the gender-distance interaction evolves over time by estimating the model separately for candidates still on the job market in their second, third, and fourth years after PhD qualification. While the main analysis focuses on first-year applicants to avoid selection bias from lower-performing candidates who remain on the market longer, this extended analysis allows me to assess whether gendered spatial frictions persist beyond initial market entry. Across all subsequent years, the gender-distance interaction remains negative and statistically significant, though its magnitude gradually declines. This suggests that spatial constraints are most binding for women at the start of their academic careers, but continue to shape application behavior even in later years.

Heterogeneity by Fields Finally, Table D17 explores whether the gender-distance interaction varies across broad disciplinary categories. The interaction is negative and statistically significant in STEM (Panel C) and Social Sciences (Panel D), where job markets are more dispersed and geographic mobility expectations higher. In contrast, the coefficients are smaller and not statistically significant in Biology and Humanities (Panels A and B), possibly due to tighter geographic clustering of job postings and smaller sample sizes. These differences point to important field-specific variation in how spatial constraints manifest across the academic labor market.

4.4 Results: Individual-level Application Behavior

To complement the dyadic analysis, I examine gender differences in the number of applications submitted, distinguishing between nearby (within 100 km) and distant (over 100 km) job opportunities. The goal is to test whether the responsiveness to local versus distant job market conditions varies by gender. I estimate Equation (2), where the outcome is the log number of applications (plus one) submitted by each candidate to jobs in either distance *type*:

$$Y_{it}^{type} = \beta_1 Female_i + \beta_2 Offers_{tf}^{type} + \beta_3 Female_i \times Offers_{tf}^{type} + X_i'\gamma + \delta_{tf} + \mu_{u(i)f} + \varepsilon_{it} \quad (2)$$

where Y_{it} denotes the number of applications that candidate i submits in year t to jobs of type *type* (either nearby or distant). The term $Offers_{tf}^{type}$ captures the number of job openings available in field f and year t for each category of distance (either under 100km or over 100km). The interaction term tests whether the responsiveness to job market thickness differs by gender. The vector X_i includes candidate-level controls for age, publication record, supervisor productivity, and supervisor gender. The model includes fixed effects for field-by-year (δ_{tf}) and for PhD institution-by-field ($\mu_{u(i)f}$), thereby accounting for both time-varying discipline-specific shocks and institutional heterogeneity in PhD institution.

Table 2: Gender Differences in Application Patterns by Distance to Job Offers

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.0212***	-0.0231***	-0.00912**	-0.0267***	-0.0334***	-0.0319***
	(0.00403)	(0.00404)	(0.00413)	(0.00870)	(0.00870)	(0.00900)
Near offers	0.0280***	0.0281***	0.0300***			
	(0.000789)	(0.000791)	(0.00124)			
Female × Near offers	0.00453***	0.00460***	0.00132			
	(0.000968)	(0.000966)	(0.00101)			
Far offers				0.0172***	0.0170***	0.0106***
				(0.000736)	(0.000733)	(0.00155)
Female × Far offers				-0.000145	-0.000137	-0.000276
				(0.000352)	(0.000350)	(0.000358)
Adj R^2	0.33	0.33	0.38	0.32	0.33	0.35
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	68258	68258	67617	68258	68258	67617

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2 reports the results. Columns (1)-(3) focus on applications to nearby jobs. Across all specifications, women submit fewer local applications than men: the female coefficient is negative and statistically significant, ranging from -0.021 to -0.009. This implies that, conditional on observables, women submit about 1–2% fewer local applications on average. However, the gap becomes smaller and loses significance in column (3), which includes PhD institution fixed effects, suggesting that institutional sorting partly explains the difference.

Importantly, the interaction term between Female and Near Offers is positive and significant in columns (1) and (2), indicating that women are more responsive to increases in the number of local job openings. In other words, women apply less overall, but are more elastic to local market conditions. turn to applications to distant jobs. Here, the gender gap is more pronounced and robust: women submit significantly fewer distant applications across all specifications (around -0.03), and the Female × Far Offers interaction is small and statistically insignificant. This suggests that female candidates are less responsive to increases in distant job availability, consistent with greater mobility constraints.

Overall, these individual-level results reinforce the findings from the dyadic analysis.

Female candidates apply to fewer jobs overall, and this gap is especially salient for distant positions. Moreover, women exhibit greater responsiveness to variation in nearby job openings but not to distant ones. These patterns support the interpretation that spatial frictions are more binding for female candidates, leading to differential application behavior even after controlling for research productivity, supervisor quality, and field-level opportunity structures.

4.5 Robustness Checks: Individual-level Results

Table D19 re-estimates the Table 1 using Poisson Pseudo-Maximum Likelihood (PPML) rather than log-linear models. The estimates remain stable in magnitude and direction, confirming that the findings are not driven by functional form assumptions.

In Tables D20–D23 I implement the same robustness checks symmetric to those shown in Table D12 to D16. Results remain consistent throughout, and are in line with the patterns observed of the dyadic approach.

Tables D24–D27 examine heterogeneity across disciplines. While the interaction between gender and geographic distance is consistently negative and significant in the pooled regressions, the coefficients are not statistically significant within individual disciplines. This likely reflects a combination of reduced statistical power and potential heterogeneity across fields. Importantly, the point estimates are generally in the same direction across disciplines, suggesting that the absence of significance may not reflect an absence of effect.

Finally, Tables D28 and D29 use alternative measures of geographic proximity, redefining "local" markets at the city and region level, respectively. The interaction between gender and local job density remains positive and significant in both cases, further supporting the interpretation that women are more sensitive to spatial constraints in their application behavior.

Taken together, these results provide robust evidence that spatial distance significantly discourages job applications, and that this deterrent effect is stronger for women. Gendered spatial constraints in the job search process persist even after conditioning on academic productivity, career stage, supervisor characteristics, and fixed effects at the candidate, job, institutional, and field level.

5 From Applications to Securing a Junior Permanent Position

The previous section documented substantial gender differences in application behavior, driven by spatial frictions. But applying is only the first step in the academic job market. In this section, I investigate how application behavior translates into hiring outcomes: are women and men equally likely to secure a position, conditional on how many jobs they apply to?

A key challenge in answering this question is that the number of applications is endogenous: more motivated or better-connected candidates may apply more and be more likely to succeed for unobserved reasons. To address this, I adopt a two-stage least squares (2SLS) strategy, using an instrument that captures exogenous variation in the spatial structure of the job market.

5.1 Application Data

@@ depends on the datastructure

This section presents descriptive statistics on the dataset used to study how geographic mobility constraints affect gender disparities in academic hiring. Table 5 summarizes the data at three levels: applications, job offers, and applicants.

In Panel A - Application-Level Statistics, the dataset contains over 180,000 candidate–job dyads. Around 44% of the applications are submitted by women. The average geographic distance between a candidate's PhD institution and the hiring university is approximately 325 kilometers. However, only 2.4% of applications result in a successful hire, reflecting the highly competitive nature of the academic job market. Panel B is at the job offer-level. Among job openings that received at least one application from newly qualified candidates, the within-sample success rate is 23%. When considering the full set of academic positions between 2009 and 2021, 94% are eventually filled. This gap reflects that many hires come from outside the sample-for instance, from earlier cohorts or international pools. On average, each job offer attracts 134 candidates, underscoring the intensity of competition for permanent academic positions in France. The Panel C includes about 30,750 unique candidates. Roughly 45% are women, but only 14.6% secure a permanent position during the observation period. On average, candidates submit about six applications, are in their early 30s (mean age: 33.5), and are about 2.5 years post-PhD. Applicants have published approximately 3 articles, with an average AIS (Article Influence Score) of 2.75. About 30% of candidates had a female PhD supervisor.

While Le Barbanchon et al. (2020) examine French jobseekers in the private sector using administrative job search data, this study focuses on a highly educated and specialized academic population. Compared to their setting, the academic market studied here features far fewer successful matches (2.4% vs. around 8% in their baseline), much longer job search durations, and more concentrated hiring processes. However, the structure is similar in terms of centralization and national coverage, enabling direct comparison of geographic mobility frictions across labor markets.

5.2 Labor Market Instrument: Job Offer Average Distance

Job Offer Average Distance Index. To proxy spatial constraints on job search, I construct a candidate-level Offer *Average Distance Index*, which measures the average geographic spread of job openings relative to each candidate's PhD institution in a specific discipline and year. Formally, it is defined as:

$$\text{Av. Distance}_i = \frac{\sum_j \text{Distance}_{ij}}{\sum_j \text{Number of Positions}_j}, \tag{3}$$

where the numerator sums the great-circle distances between candidate i's PhD institution and each job posting j in their discipline and year, and the denominator is the total number of such positions. Intuitively, a higher value reflects a more spatially dispersed job market.

Instrument Validity. The identifying variation comes from differences in job market structure across time, space, and disciplines - variation that is plausibly exogenous due to the timing of the hiring cycle in French academia.

Candidates must obtain a national qualification ("qualification CNU") before entering the job market. This occurs in the fall of year $t-1$. The list of job openings, however, is released only in the spring of year t, once candidates have already completed their PhD and qualification process. As a result, applicants cannot anticipate where or how many jobs will open when they prepare their applications.

Moreover, qualification is discipline-specific, and candidates apply to jobs in their qualified discipline. These institutional features make the spatial structure of the market - including the average distance of job postings - a plausibly exogenous shock to individual application decisions.

5.3 Estimation Model

To estimate the causal effect of applying to more jobs on the probability of securing a permanent academic position, I adopt a two-stage least squares (2SLS) strategy. The challenge lies in the endogeneity of application behavior: candidates who apply to more jobs may differ systematically in unobserved ways (e.g., motivation, mobility constraints). To address this, I instrument log applications using a *Average Distance Index*, which captures the spatial dispersion of job opportunities available to a candidate in their discipline and year. The first-stage equation is given by:

$$\begin{aligned}\ln(\text{Applications}_{ift}+1) = \pi_0 + \pi_1\,\text{Av. Distance}_{ift} + \pi_2\,\mathit{Female}_i \times \text{Av. Distance}_{ift} \\ + X_i'\lambda + \delta_f + \delta_t + u_i\end{aligned} \tag{4}$$

The second-stage equation estimates the impact of (instrumented) application intensity on hiring outcomes:

$$\begin{aligned}\text{Success}_{ift} = \beta_0 + \beta_1 \ln(\widehat{\text{Applications}}_{ift}+1) + \beta_2\,\mathit{Female}_i \\ + \beta_3\,\mathit{Female}_i \times \ln(\widehat{\text{Applications}}_{ift}+1) + X_i'\gamma + \delta_{ft} + \delta_{uf} + \varepsilon_i\end{aligned} \tag{5}$$

In both equations, the subscript i indexes candidates, f denotes the academic discipline, and t the year of first application. The dependent variable in the second stage, Success_{ift}, is a binary indicator equal to 1 if candidate i secures a junior permanent position in field f.

The main instrument, $\text{Av. Distance}_{ift}$, is the *Av. Distance Index*, defined as the average great-circle distance between candidate i's PhD institution and all job openings in discipline f and year t. Higher values indicate a more spatially dispersed market, which increases search costs and reduces the feasibility of applying to multiple jobs.

The control vector X_i includes candidate-level characteristics such as age, research productivity (quantity and quality), and supervisor productivity. The specification includes fixed effects for discipline-by-year (δ_{ft}), which capture time-varying shocks to academic hiring within fields, and PhD university-by-field fixed effects (δ_{uf}), which account for persistent institutional differences in candidate quality or preparation.

5.4 Results

Table 3 presents OLS and 2SLS estimates of the effect of application intensity on the probability of securing a permanent academic position. The key endogenous variable is the log number of applications submitted, instrumented using the *Average Distance Index*-a measure of how spatially dispersed job opportunities are for each candidate's field and year.

Columns (1) and (2) report standard OLS regressions. In column (1), the coefficient on $\ln(\text{Applications}+1)$ is 0.098 and highly significant, indicating that a 10% increase

Table 3: Instrumental Variables Estimates of the Effect of Applications on Success

	OLS (1)	OLS (2)	First stage (3)	IV (4)	IV (5)	IV (6)	Reduced form (7)
Dependent variable:	*Success*	*Success*	*ln(Apps+1)*	*Success*	*Success*	*Success*	*Success*
Female	-0.00908***	-0.00489	0.0348	0.271	0.047	0.022	0.0116
	(0.00305)	(0.00318)	(0.0287)	(0.233)	(0.057)	(0.048)	(0.00895)
ln(Apps+1)	0.0978***	0.0855***		0.332***	0.340***	0.281***	
	(0.00151)	(0.00167)		(0.085)	(0.104)	(0.080)	
Female × ln(Apps+1)	0.00716***	0.00346		-0.088	-0.044	-0.019	
	(0.00228)	(0.00230)		(0.068)	(0.063)	(0.054)	
Av. Distance			-0.000346***				-0.0000986***
			(0.0000874)				(0.0000273)
Female × Av. Distance			-0.000211***				-0.0000470**
			(0.0000752)				(0.0000235)
Controls		yes	yes	yes	yes	yes	yes
Fields X Year FE		yes	yes	yes		yes	yes
Fields X Univ PhD FE		yes	yes		yes	yes	yes
F-stat (excluded instruments)			13.02				
Observations	51544	51391	51391	51544	51544	51544	51391

Notes: *Success* is a binary indicator for securing a permanent position. *ln(Apps + 1)* is instrumented using the average distance between available jobs and the candidate's PhD location, as well as its interaction with a female dummy. All regressions control for age, publication metrics, and supervisor gender. Column (3) reports the first stage; column (7) reports the reduced form. The F-statistic reported in column (3) tests the joint significance of the excluded instruments. First-stage results for the IV estimations in columns (4) and (5) are available in Appendix Table @@. Robust standard errors in parentheses. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in applications is associated with a 0.98 percentage point increase in the probability of success. Column (2), which adds controls for age, publications, supervisor quality, and fixed effects, shows a slightly smaller effect of 0.085. The interaction with the Female dummy is positive in both cases (0.007 and 0.003, respectively), but the coefficient loses significance when controls are added. These estimates suggest that women may benefit slightly more from increased application intensity, but the differences are modest and potentially confounded by selection.

Column (3) presents the first stage of the 2SLS estimation. The *Av. Distance Index* (average distance to jobs in a candidate's field and year) is strongly and negatively associated with the number of applications submitted. A one-kilometer increase in average distance reduces log applications by 0.00035, significant at the 1% level. The interaction with gender is also negative and significant, suggesting that female candidates are especially discouraged by dispersed markets. The F-statistic for the excluded instruments is 13.02, above conventional thresholds. Columns (4)-(6) report the second-stage IV estimates. Column (4), which includes basic controls and field-by-year fixed effects, yields a large and significant coefficient of 0.332 on instrumented log applications: a 10% increase in applications causally raises the probability of obtaining a permanent job by 3.3 percentage points. Column (5), which adds PhD university-by-field fixed effects, produces a similar estimate (0.340), reinforcing the robustness of the result.

Column (6) adds the full set of controls and fixed effects and yields a slightly smaller coefficient (0.281), still statistically significant at the 1

In contrast to the OLS results, the interaction between female and log applications is now negative and statistically insignificant across all IV specifications. This suggests that once selection bias is accounted for, there is no evidence that women derive higher marginal returns from applying to more jobs. In fact, if anything, the point estimates suggest slightly lower returns.

Column (7) reports the reduced-form regression of success on the Job Av. Distance Index. The coefficient is negative and statistically significant, confirming that candidates in more spatially fragmented markets-who submit fewer applications-are less likely to secure a position.

Robustness. @@

Taken together, the IV estimates support a causal interpretation of the relationship between application behavior and hiring success. Candidates who face more spatially accessible markets apply more and are significantly more likely to secure a job. Moreover, gender differences in marginal returns to applying disappear once endogeneity is addressed.

6 Conclusion

This paper investigates how geographic constraints shape application behavior and hiring outcomes in the French academic job market, with a focus on gender disparities. Using administrative data covering the universe of PhD graduates and job openings from 2009 to 2021, I construct a novel dyadic dataset linking candidates to all job opportunities in their field. This framework enables me to trace how spatial frictions - particularly distance from the PhD institution - affect both the decision to apply and the likelihood of securing a permanent academic position.

I document three main findings. First, distance significantly reduces the probability of applying to a given position, and this effect is more pronounced for female candidates. Women are less likely to apply to distant jobs and more responsive to local job availability, consistent with greater spatial constraints. Second, using an instrumental variable strategy based on exogenous variation in the geographic dispersion of job offers, I show that applying to more positions causally increases the probability of success. This suggests that spatial frictions - by limiting application intensity - translate directly into lower hiring rates. Third, I find that once this constraint is accounted for, there is no evidence that women benefit from higher marginal returns to applying. Instead, the gender gap in outcomes arises largely from differential exposure to mobility barriers.

Taken together, these results highlight the importance of market structure in shaping early academic careers. Because job openings are unevenly distributed across space and disciplines, candidates with limited geographic mobility - disproportionately women - face systematically lower chances of success.

References

Antecol, Heather, Kelly Bedard, and Jenna Stearns, "Equal but Inequitable: Who Benefits from Gender-Neutral Tenure Clock Stopping Policies?," *American Economic Review*, September 2018, *108* (9), 2420–41.

Bagues, Manuel, Mauro Sylos-Labini, and Natalia Zinovyeva, "Does the Gender Composition of Scientific Committees Matter?," *American Economic Review*, April 2017, *107* (4), 1207–38.

_ , _ , and _ , "Does the Gender Composition of Scientific Committees Matter?," *American Economic Review*, April 2017, *107* (4), 1207–38.

Barbanchon, Thomas Le, Roland Rathelot, and Alexandra Roulet, "Gender Differences in Job Search: Trading off Commute against Wage*," *The Quarterly Journal of Economics*, 10 2020, *136* (1), 381–426.

Benveniste, Stéphane, "Like Father, Like Child: Intergenerational Mobility in the French Grandes Écoles throughout the 20 th Century," *WP*, 2023.

Bisantis, Aliénor, "Gender Disparities in French Academic Careers: A Multi-Stage Analysis of Selection," 2025. Manuscript in preparation. Available upon request.

Bosquet, Clément, Pierre-Philippe Combes, and Cecilia García-Peñalosa, "Gender and promotions: Evidence from academic economists in France," *The Scandinavian Journal of Economics*, 2019, *121* (3), 1020–1053.

Card, David, Stefano DellaVigna, Patricia Funk, and Nagore Iriberri, "Gender differences in peer recognition by economists," *Econometrica*, 2022, *90* (5), 1937–1971.

Corsini, Alberto, Michele Pezzoni, and Fabiana Visentin, "What makes a productive Ph.D. student?," *Research Policy*, 2022, *51* (10), 104561.

Diamond, Rebecca, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, March 2016, *106* (3), 479–524.

Dossi, Gaia, "Race and Science," *WP*, 2024, (4).

Ductor, Lorenzo, Sanjeev Goyal, and Anja Prummer, "Gender and collaboration," *Review of Economics and Statistics*, 2023, *105* (6), 1366–1378.

Galván, Estefanía and Victoria Tenenbaum, "Gender Gaps in Academia: The Role of Children," *WP*, 2024.

Gaule, Patrick and Mario Piacentini, "An advisor like me? Advisor gender and post-graduate careers in science," *Research Policy*, 2018, *47* (4), 805–813.

Ginther, Donna and Shulamit Kahn, "Women in Economics: Moving Up or Falling Off the Academic Career Ladder?," *Journal of Economic Perspectives*, 02 2004, *18*, 193–214.

Hengel, Erin, "Publishing while female: Are women held to higher standards? Evidence from peer review," *The Economic Journal*, 2022, *132* (648), 2951–2991.

J., Williams W. M. Ceci S., "Understanding current causes of women's underrepresentation in science.," *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 2011.

J., Gates A. J. Sinatra R. Barabási A.-L. Huang, ""Historical Comparison of Gender Inequality in Scientific Careers across Countries and Disciplines."," *Proceedings of the National Academy of Sciences*, 2020.

J., Ginther D. K. Kahn S. Williams-W. M. Ceci S., "Women in Academic Science: A Changing Landscape," *Psychological science in the public interest : a journal of the American Psychological Society*, 2014.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard, "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, October 2019, *11* (4), 181–209.

_ , _ , Mathilde Muñoz, and Stefanie Stantcheva, "Taxation and Migration: Evidence and Policy Implications," *Journal of Economic Perspectives*, May 2020, *34* (2), 119–42.

Lassen, Anne Sophie and Ria Ivandić, "Parenthood and Academic Career Trajectories," *AEA Papers and Proceedings*, May 2024, *114*, 238–42.

Lerchenmueller, Marc J. and Olav Sorenson, "The gender gap in early career transitions in the life sciences," *Research Policy*, 2018, *47* (6), 1007–1017.

Mayer, Thierry and Soledad Zignago, "Notes on CEPIIâ€™s distances measures: The GeoDist database," Working Papers 2011-25, CEPII research center Dec 2011.

Meyer, Meredith, Andrei Cimpian, and Sarah-Jane Leslie, "Women are underrepresented in fields where success is believed to require brilliance," *Frontiers in Psychology*, 2015, *6*.

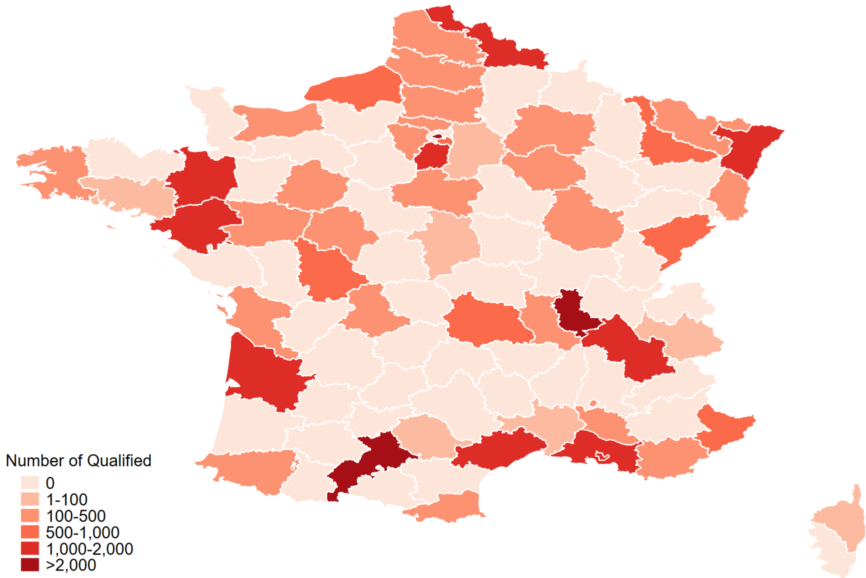
Patsali, Sofia, Michele Pezzoni, and Fabiana Visentin, "Research independence: drivers and impact on PhD students' careers," *Studies in Higher Education*, 02 2024, pp. 1–24.

Sarsons, Heather, "Recognition for Group Work: Gender Differences in Academia," *American Economic Review*, May 2017, *107* (5), 141–145.

Wu, Alice H., "Gendered Language on the Economics Job Market Rumors Forum," *AEA Papers and Proceedings*, May 2018, *108*, 175–79.

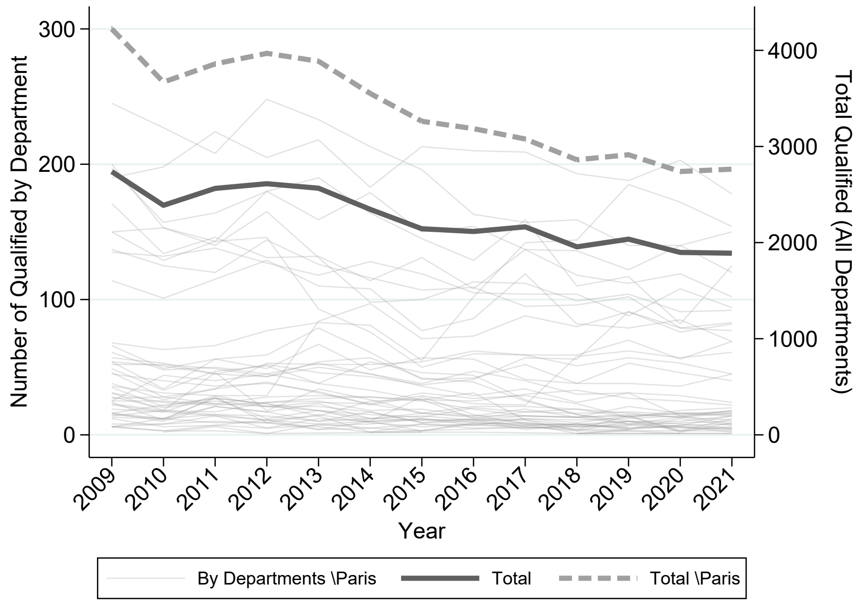
Figures

Figure 3: Cumulative Number of Qualified Candidates by Department of PhD (2009–2021)



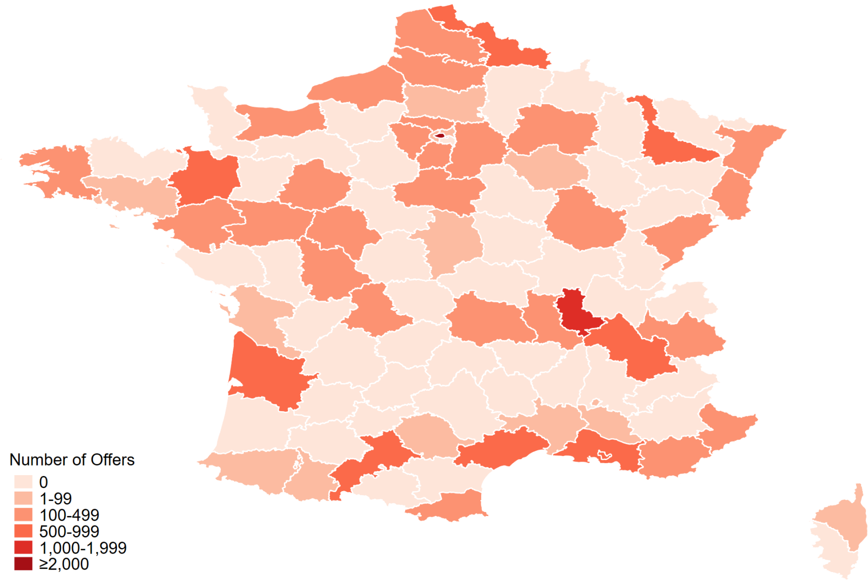
Notes: This map shows the total number of candidates qualified for junior permanent academic positions (*Maître de Conférence*) between 2009 and 2021, based on the city location of their PhD institution. Values are aggregated at the departmental level (96 mainland French departments). Departments with darker shading indicate higher numbers of qualified candidates. The spatial distribution is highly concentrated, with Paris (département 75) alone accounting for over 14,500 qualifiers - nearly 30% of the national total. 30 rural or peripheral departments recorded zero qualifiers over the same period.

Figure 4: Annual Number of Qualified Candidates by Department (2009–2021)



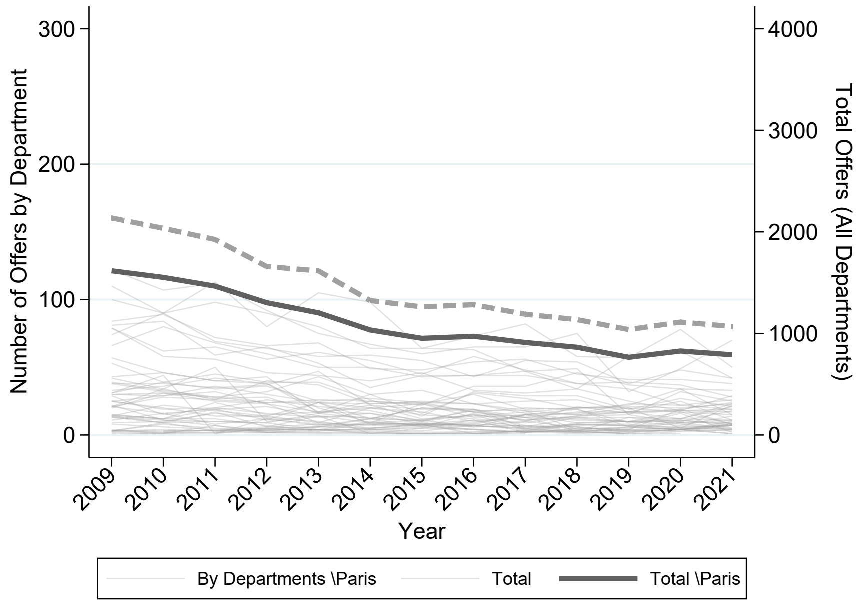
Notes: The figure shows the annual number of candidates qualified to apply for junior permanent academic positions (*Maître de Conférence*) from 2009 to 2021, by *department* of PhD graduation's city. Department-level trends (left y-axis) exclude Paris to improve readability. Two national totals are shown on the right y-axis: the dashed line includes Paris, while the solid line excludes it. Paris is excluded from the department lines due to its much larger volume (over 14,500 qualifiers during the period), which would otherwise compress variation across other departments.

Figure 5: Cumulative Number of Permanent Academic Job Offers by Department (2009–2021)



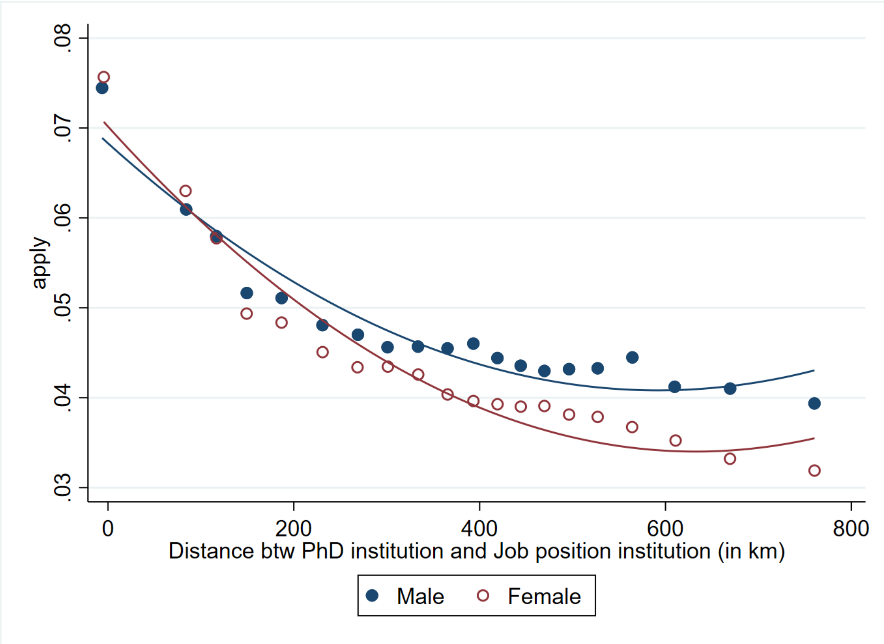
Notes: This map shows the cumulative number of junior permanent academic (*Maître de Conférence*) job offers between 2009 and 2021, aggregated by the *department* of the hiring institution. The color scale is consistent with Figure 3 (qualified candidates), allowing for visual comparison. Paris (département 75) had the highest number of positions (4,529), while more than 20 departments recorded zero offers during this period.

Figure 6: Annual Number of Permanent Academic Job Offers by Department (2009–2021)



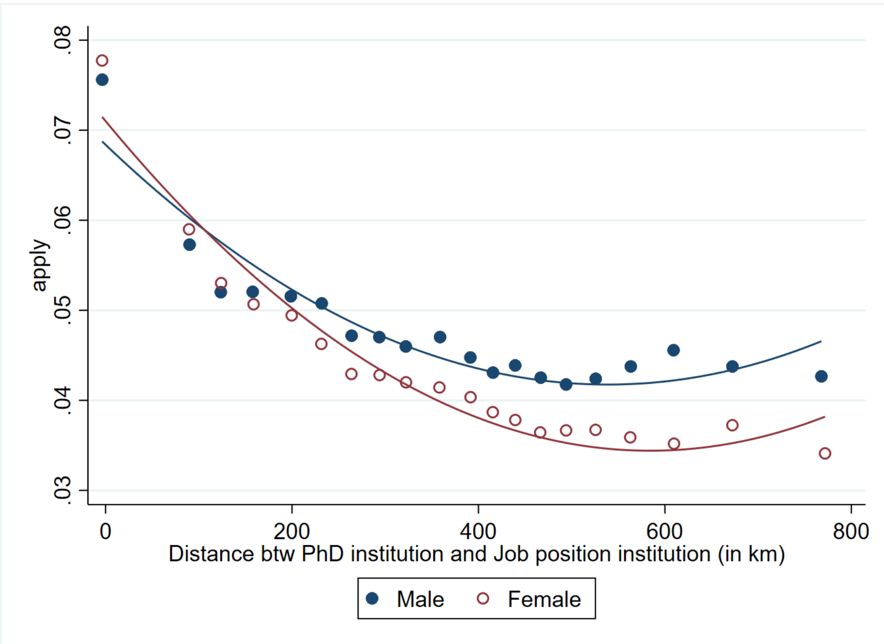
Notes: This figure shows the number of junior permanent academic (*Maître de Conférence*) job offers from 2009 to 2021. Department-level trends are plotted on the left y-axis (excluding Paris for readability). The right y-axis displays national totals: the dashed line includes Paris, while the solid line excludes it. Paris is excluded from department-level lines due to its large scale, which would otherwise compress variation across other departments.

Figure 7: Predicted Number of Applications by Distance to Job Offers by Gender



Notes: The figure presents a binned scatterplot of the application rate versus the distance from the candidate's PhD institution and the position institution, shown separately by gender. The application rate and distance are residualized controlling for age, publication metrics, supervisor characteristics, and PhD institution × Field × year fixed effects.

Figure 8: Predicted Number of Applications by Distance to Job Offers by Gender



Notes: The figure presents a binned scatterplot of the application rate versus the distance from the candidate's PhD institution and the position institution, shown separately by gender. The application rate and distance are residualized controlling for age, publication metrics, supervisor characteristics, and Job position's institution × Field × year fixed effects.

Tables

Table 4: Descriptive Statistics

	Male			*Female*		
Variable	N	Mean	SD	N	Mean	SD
Apply Position	28,822	0.51	0.50	22,722	0.53	0.50
Number Applications	28,822	3.55	7.44	22,722	3.57	7.15
Securing Position	28,822	0.09	0.28	22,722	0.09	0.28
Age	28,822	33.10	5.78	22,722	33.71	6.01
Time since PhD	28,822	2.89	2.85	22,722	2.82	2.78
Publish	28,822	0.64	0.48	22,722	0.49	0.50
Number Publications	28,822	5.37	18.01	22,722	2.90	11.70
Total AIS	28,822	5.85	34.76	22,722	3.29	19.44
Female Supervisor	28,822	0.24	0.43	22,722	0.36	0.48
Total AIS Supervisor	28,822	0.07	0.96	22,722	0.06	0.89
Disciplines						
Biological Science	24,730	0.07	0.26	19,241	0.12	0.32
Chemical Science	24,730	0.05	0.22	19,241	0.04	0.21
Computer Science	24,730	0.11	0.31	19,241	0.04	0.21
Earth Science	24,730	0.04	0.20	19,241	0.05	0.22
Economics	24,730	0.03	0.18	19,241	0.03	0.17
Engineering	24,730	0.13	0.34	19,241	0.06	0.25
Humanities	24,730	0.18	0.38	19,241	0.31	0.46
Law and Political Science	24,730	0.05	0.22	19,241	0.06	0.23
Literature	24,730	0.05	0.21	19,241	0.12	0.32
Management Sciences	24,730	0.03	0.17	19,241	0.05	0.21
Mathematics	24,730	0.08	0.27	19,241	0.03	0.18
Philosophy and Theology	24,730	0.03	0.16	19,241	0.02	0.15
Physical Science	24,730	0.15	0.36	19,241	0.07	0.25

Notes: This table presents statistics for the key variables in the paper and the different disciplines of the qualified PhD graduates at their first year of application, if they were interested in applying for at least one position.

Table 5: Summary Statistics of Application Dataset

	Mean (1)	Std. dev. (2)	Obs (3)
Panel A: Application level			183,238
Secure position	0.024	0.155	
Female applicants	0.442	0.497	
Distance (km)	325.117	233.884	
Panel B: Job offers level			
Secured position (Sample)[a]	0.234	0.423	18,785
Secured position (Total)[b]	0.939	0.239	22,688
Number applicants per offer (Total)[b]	133.916	116.984	
Panel C: Applicant level			30,750
Female	0.452	0.498	
Secure position	0.146	0.353	
Number applications	5.959	8.261	
Age	33.475	5.808	
Time since PhD	2.521	2.605	
Number Publications	3.187	11.718	
Total AIS	2.754	20.76	
Female Supervisor	0.295	0.456	
Total AIS Supervisor	0.021	0.765	

Notes: This table reports summary statistics on qualified candidate's application for junior permanent positions offers. In Panel A, I report statistics at the application level. In Panel B, I collapse the data set at the offer level. in Panel C, I collapse the data set at the applicant/qualified level.

[a]Represents the success rate in the sample of PhD graduates from France qualified and applying for at least one position the first year of qualification - the sample used in the estimation.

[b]Represents the total sample of job offers between 2009 and 2021 and the success rate among all candidates

A Appendix

Fields	Section	Label (fr/eng)
Law and Political Science	01	Droit privé et sciences criminelles - Private law and criminal sciences
	02	Droit public - Public law
	03	Histoire du droit et des institutions - History of law and institutions
	04	Science politique - Political Science
Economics	05	Sciences économiques
Management	06	Sciences de gestion et du management
Literature	07	Sciences du langage - Language sciences
	08	Langues et littératures anciennes - Ancient languages and literature
	09	Langue et littérature française - French language and literature
	10	Littératures comparées - Comparative literature
	11	Études anglophones - English-language studies
	12	Études germaniques et scandinaves - Germanic and Scandinavian Studies
	13	Études slaves et baltes - Slavic and Baltic Studies
	14	Études romanes - Romance languages and literature
	15	Langues, littératures et cultures africaines, asiatiques et d'autres aires linguistiques - Languages, literatures and cultures of Africa, Asia and other linguistic areas
	73	Cultures et langues régionales - Regional cultures and languages
Humanities	16	Psychologie et ergonomie - Psychology and ergonomics
	18	Architecture (ses théories et ses pratiques), arts appliqués, arts plastiques, arts du spectacle, épistémologie des enseignements artistiques, esthétique, musicologie, musique, sciences de l'art - Arts
	19	Sociologie, démographie - Sociology, demography
	20	Ethnologie, préhistoire, anthropologie biologique - Biological anthropology, ethnology, prehistory
	21	Histoire, civilisations, archéologie et art des mondes anciens et médiévaux - History, civilization: archaeology, art of ancient worlds
	22	Histoire et civilisations : histoire des mondes modernes, histoire du monde contemporain ; de l'art ; de la musique - History, civilizations: history of modern worlds
	23	Géographie physique, humaine, économique et régionale - Physical, human, economic and regional geography
	24	Aménagement de l'espace, urbanisme - Spatial planning and urban development
	70	Sciences de l'éducation et de la formation - Education sciences
	71	Sciences de l'information et de la communication - Information and communication sciences
	72	Epistémologie, histoire des sciences et des techniques - Epistemology, history of science and technology
Mathematics	25	Mathématiques - Mathematics
	26	Mathématiques appliquées et applications des mathématiques - Applied mathematics and mathematical applications
Computer Science	27	Informatique - Computer science
Physical Science	28	Milieux denses et matériaux - Dense media and materials
	29	Constituants élémentaires - Elementary constituents
	30	Milieux dilués et optique - Diluted media and optics
Chemical Science	31	Chimie théorique, physique, analytique - Theoretical, physical and analytical chemistry
	32	Chimie organique, minérale, industrielle - Organic, inorganic and industrial chemistry
	33	Chimie des matériaux - Materials chemistry
Earth Science	34	Astronomie, astrophysique - Astronomy, astrophysics
	35	Structure et évolution de la terre et des autres planètes - Structure and evolution of the Earth and other planets
	36	Terre solide : géodynamique des enveloppes supérieure, paléobiosphère - Solid Earth: geodynamics of the upper envelope
	37	Enveloppes fluides du système Terre et autres planètes - Fluid envelopes of the Earth system and other planets
Engineering	60	Mécanique, génie mécanique, génie civil - Mechanical engineering, civil engineering
	61	Génie informatique, automatique et traitement du signal - Computer engineering, automation and signal processing
	62	Energétique, génie des procédés - Energy and process engineering
	63	Génie électrique, électronique, photonique et systèmes - Electrical engineering, electronics, photonics and systems
Biological Science	64	Biochimie et biologie moléculaire - Biochemistry and molecular biology
	65	Biologie cellulaire - Cell Biology
	66	Physiologie - Physiology
	67	Biologie des populations et écologie - Population biology and ecology
	68	Biologie des organismes - Organismal biology
	69	Neurosciences - Neuroscience
Philosophy and Theology	76	Théologie catholique - Catholic theology
	77	Théologie protestante - Protestant theology
	17	Philosophie - Philosophy
Medical Science	85	Personnels enseignants-chercheurs de pharmacie en sciences physico-chimiques et ingénierie appliquée à la santé - Engineering applied to health
	86	Personnels enseignants-chercheurs de pharmacie en sciences du médicament et des autres produits de santé - Sciences of drugs and other health products
	87	Personnels enseignants-chercheurs de pharmacie en sciences biologiques, fondamentales et cliniques - Biological, fundamental and clinical sciences
	90	Maïeutique - Maieutics
	91	Personnels enseignants-chercheurs des disciplines des sciences de la rééducation et de réadaptation - Rehabilitation sciences
	92	Personnels enseignants-chercheurs des disciplines des sciences infirmières - Nursing
	74	Sciences et techniques des activités physiques et sportives - Sciences and techniques of physical activities and sports

Table A1: CNU Sections and Labels

B Descriptive Statistics

B.1 Success Sample

Table B2: Descriptive Statistics - Success Sample

	Male			*Female*		
Variable	N	Mean	SD	N	Mean	SD
Number Applications	2,503	11.15	12.63	1,981	11.94	12.57
Age	2,503	31.79	4.87	1,981	32.51	5.18
Time since PhD	2,503	1.99	2.12	1,981	1.89	1.99
Publish	2,503	0.60	0.49	1,981	0.44	0.50
Number Publications	2,503	4.16	9.22	1,981	1.95	4.71
Total AIS	2,503	3.25	26.08	1,981	1.59	7.49
Female Supervisor	2,503	0.23	0.42	1,981	0.37	0.48
Total AIS Supervisor	2,503	0.03	0.76	1,981	0.02	0.58
Disciplines						
Biological Science	2,503	0.03	0.18	1,981	0.04	0.19
Chemical Science	2,503	0.03	0.17	1,981	0.01	0.12
Computer Science	2,503	0.11	0.31	1,981	0.04	0.19
Earth Science	2,503	0.02	0.13	1,981	0.01	0.11
Economics	2,503	0.06	0.24	1,981	0.07	0.25
Engineering	2,503	0.13	0.33	1,981	0.05	0.22
Humanities	2,503	0.15	0.36	1,981	0.26	0.44
Law and Political Science	2,503	0.14	0.35	1,981	0.16	0.37
Literature	2,503	0.06	0.23	1,981	0.13	0.34
Management Sciences	2,503	0.09	0.28	1,981	0.14	0.34
Mathematics	2,503	0.08	0.28	1,981	0.04	0.20
Philosophy and Theology	2,503	0.01	0.10	1,981	0.01	0.11
Physical Science	2,503	0.09	0.28	1,981	0.03	0.18

Notes: This table presents statistics for the key variables in the paper and the different disciplines of the qualified PhD graduates at their first year of application, if they were interested in applying for at least one position.

B.2 Descriptive Statistics by Fields

Table B3: Descriptive Statistics - Biological & Earth Sciences

	Male			*Female*		
Variable	N	Mean	SD	N	Mean	SD
Apply Position	3,521	0.30	0.46	3,821	0.27	0.44
Number Applications	3,521	0.73	1.94	3,821	0.59	1.64
Securing Position	3,521	0.04	0.19	3,821	0.03	0.16
Age	3,521	32.36	3.96	3,821	31.73	3.62
Time since PhD	3,521	4.12	3.13	3,821	3.82	3.00
Publish	3,521	0.52	0.50	3,821	0.55	0.50
Number Publications	3,521	5.15	8.36	3,821	4.72	12.00
Total AIS	3,521	9.65	18.21	3,821	9.01	20.08
Female Supervisor	3,521	0.32	0.47	3,821	0.39	0.49
Total AIS Supervisor	3,521	-0.04	0.72	3,821	0.02	0.93

Notes: This table presents statistics for the key variables in the paper for the field of Biological and Earth Sciences of the qualified PhD graduates at their first year of application, if they were interested in applying for at least one position.

Table B4: Descriptive Statistics - Humanities

	Male			*Female*		
Variable	N	Mean	SD	N	Mean	SD
Apply Position	7,847	0.62	0.48	10,719	0.63	0.48
Number Applications	7,847	2.91	4.34	10,719	3.01	4.26
Securing Position	7,847	0.07	0.26	10,719	0.08	0.26
Age	7,847	37.30	6.99	10,719	36.22	6.70
Time since PhD	7,847	3.28	3.23	10,719	2.95	2.97
Publish	7,847	0.43	0.49	10,719	0.36	0.48
Number Publications	7,847	1.28	3.03	10,719	0.89	2.11
Total AIS	7,847	0.39	2.64	10,719	0.23	1.37
Female Supervisor	7,847	0.29	0.45	10,719	0.40	0.49
Total AIS Supervisor	7,847	0.04	0.97	10,719	0.03	0.79

Notes: This table presents statistics for the key variables in the paper for the field of Humanities of the qualified PhD graduates at their first year of application, if they were interested in applying for at least one position.

Table B5: Descriptive Statistics - STEM

	Male			*Female*		
Variable	N	Mean	SD	N	Mean	SD
Apply Position	14,396	0.45	0.50	5,384	0.42	0.49
Number Applications	14,396	2.59	5.60	5,384	2.21	4.90
Securing Position	14,396	0.08	0.26	5,384	0.07	0.25
Age	14,396	30.89	3.96	5,384	30.40	3.65
Time since PhD	14,396	2.52	2.51	5,384	2.29	2.20
Publish	14,396	0.84	0.37	5,384	0.80	0.40
Number Publications	14,396	8.59	24.56	5,384	6.73	20.90
Total AIS	14,396	9.06	47.93	5,384	6.87	35.27
Female Supervisor	14,396	0.20	0.40	5,384	0.29	0.45
Total AIS Supervisor	14,396	0.13	1.03	5,384	0.17	1.07

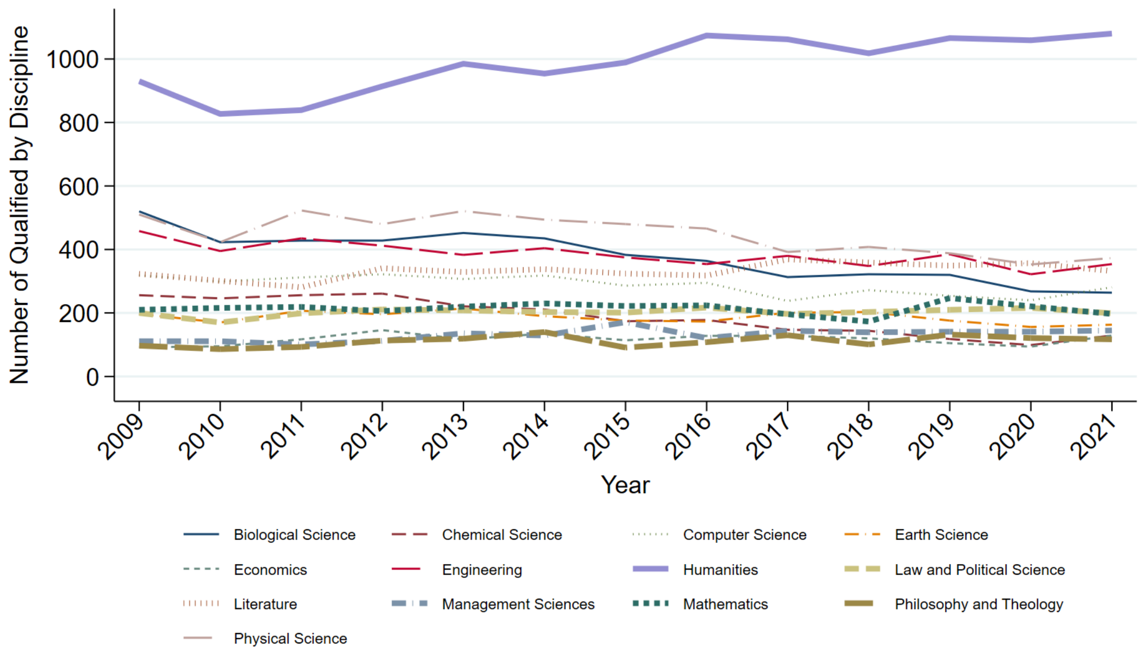
Notes: This table presents statistics for the key variables in the paper for the field of STEM of the qualified PhD graduates at their first year of application, if they were interested in applying for at least one position.

Table B6: Descriptive Statistics - Social Sciences

	Male			*Female*		
Variable	N	Mean	SD	N	Mean	SD
Apply Position	3,058	0.75	0.43	2,798	0.74	0.44
Number Applications	3,058	12.94	14.79	2,798	12.39	14.20
Securing Position	3,058	0.24	0.43	2,798	0.26	0.44
Age	3,058	33.58	5.33	2,798	33.15	5.16
Time since PhD	3,058	2.24	2.38	2,798	1.98	2.17
Publish	3,058	0.35	0.48	2,798	0.33	0.47
Number Publications	3,058	0.98	2.15	2,798	0.79	2.05
Total AIS	3,058	0.40	1.71	2,798	0.30	1.36
Female Supervisor	3,058	0.24	0.43	2,798	0.33	0.47
Total AIS Supervisor	3,058	0.02	0.77	2,798	0.05	0.76

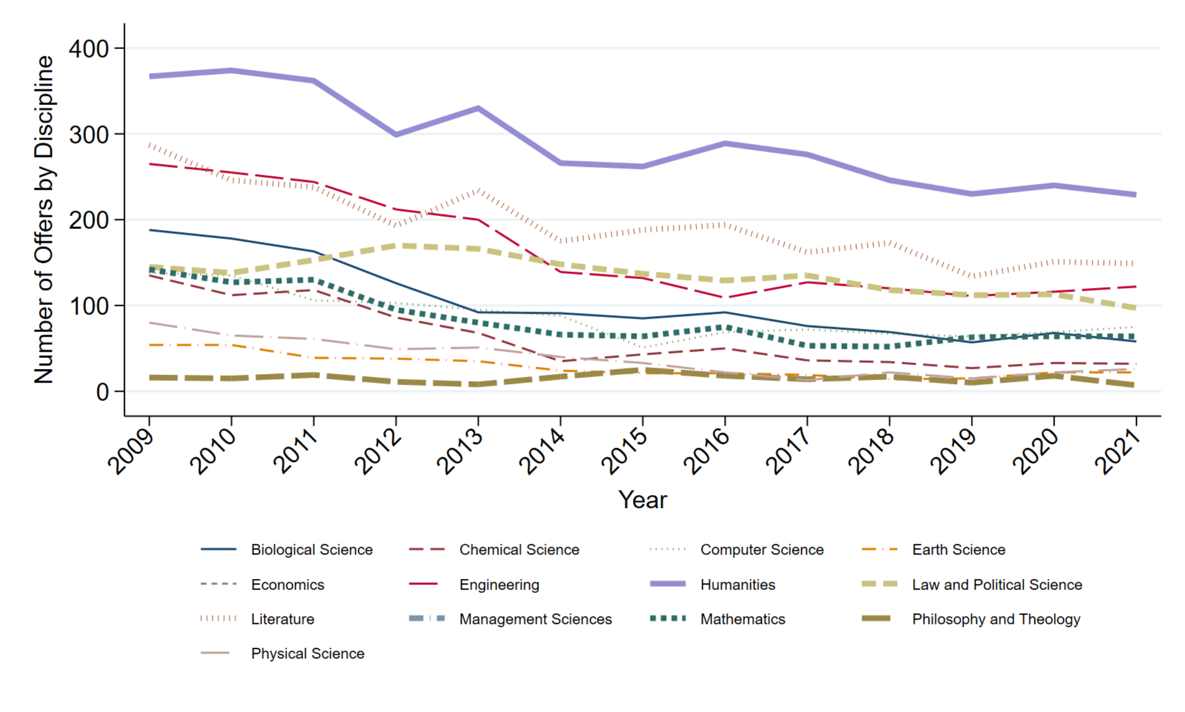
Notes: This table presents statistics for the key variables in the paper for the field of Social Sciences of the qualified PhD graduates at their first year of application, if they were interested in applying for at least one position.

Figure B1: Number of Qualified by Discipline, 2009–2021



Notes: This figure plots the annual number of qualified candidates for junior permanent position (*Maître de Conférence*), disaggregated by discipline.

Figure B2: Number of Job Offers by Discipline, 2009–2021



Notes: This figure plots the annual number of junior position (*Maître de Conférence*) offers in French public universities, disaggregated by discipline.

B.3 Decomposition Method

The progression from PhD to permanent position involves three sequential transitions: (1.a) Application for qualification after PhD (AQ), (1.b) Qualification success conditional on applying ($Q|AQ$), and (2) Secure a permanent position conditional on qualification ($JP|Q$).

The unconditional probability of securing a permanent position can be expressed as the product of these three conditional probabilities:

$$Pr(S) = Pr(AQ) \times Pr(Q|AQ) \times Pr(JP|Q) \tag{6}$$

The gender gap in this unconditional probability is:

$$\Delta Pr(S) = Pr(S;m) - Pr(S;f) \tag{7}$$

where m and f denote men and women. This can be expanded as:

$$\Delta Pr(S) = Pr(AQ; m) \times Pr(Q|AQ; m) \times Pr(JP|Q; m)$$
$$-Pr(AQ; f) \times Pr(Q|AQ; f) \times Pr(JP|Q; f) \quad (8)$$

For each stage, I decompose the contribution to the overall gender gap into application and success. For example, for the first transition (PhD to qualification), the gap between obtaining qualification and not can be decomposed as:

$$\Delta Pr(Q) = \overline{Pr(Q|AQ)} \times \Delta Pr(AQ) + \overline{Pr(AQ)} \times \Delta Pr(Q|AQ) \quad (9)$$

Where $\overline{Pr(Q|AQ)}$ and $\overline{Pr(AQ)}$ are the averAgeprobabilities across genders[12].

Similarly, I can identify the contribution of each transition to the overall gender gap. For example, the contribution of the application for the qualification stAgecan be expressed as:

$$\text{Contribution of AQ} = \overline{Pr(Q|AQ)} \times \overline{Pr(JP|Q)} \times \Delta Pr(AQ) \quad (10)$$

This approach allows me to determine whether gender gaps arise primarily from differences in application behavior or from differences in success rates, and to quantify what percentAgeof the overall gender gap is attributable to each specific transition in the academic pipeline.

Linear probability regression model:

To estimate the conditional probability of success of individual i at time t, PhD graduates from university u, in field f at each transition stage, I follow the methodology of Bosquet et al. (2019) and use a linear probability model for all probabilities. My empirical analysis considers four sequential transitions in the academic career path: (1.a) Application for qualification after PhD (AQ), (1.b) Qualification success conditional on applying ($Q|AQ$), (2) Secure a junior permanent position conditional on qualification ($JP|Q$). For an outcome O where O $\in \{AQ;\ Q|AQ;\ JP|Q\}$, I estimate:

$$\begin{aligned}\text{Pr(O)}_{ituf} = \beta_0 + \beta_1 \text{Female}_i + \beta_2 \text{Time since PhD}_{\ it} + \beta_3 \text{Time since PhD}_{\ it}^{\ 2} \\ + \beta_4 \text{Publish}_{it} + \beta_5 (\text{Publish} \times \text{Quantity})_{it} + \beta_6 (\text{Publish} \times \text{Quality})_{it} \\ + \beta_7 \text{Female Supervisor}_i + \beta_8 \text{Quality\_supervisor}_i + \alpha_{uf} + \gamma_t + \epsilon_{ituf} \quad (11)\end{aligned}$$

[12] $\overline{Pr(X)} = \frac{Pr(X;m)+Pr(X;f)}{2}$

The outcome is a function of experience since PhD graduation ($TimesincePhD_{it}$) and its square, whether individual i has at least one scientific publication appearing in the *Scopus platform* (dummy $Publish_{it}$), the cumulative number of publications at year t ($Quantity_{it}$) and the cumulative Article Influence Score (AIS) of publications at year t ($Quality_{it}$), and supervisor characteristics including whether at least one supervisor is female ($FemaleSupervisor$) and the cumulative AIS score of supervisors at the year of PhD defense of individual i ($Quality\_supervisor_i$). $Female_i$ is a dummy variable equal to 1 if the PhD graduate is female and 0 if male; β_1 measures the gender differences in probability for individuals with the same characteristics. γ_t are year fixed effects that capture time-specific trends in a non-parametric manner. α_{uf} are university-field fixed effects that control for local factors affecting PhD graduates' academic productivity, such as departments' social capital and academic quality.

C Data Thèses.fr - Detailed Procedure

We construct our dataset using data from *Theses.fr*, which provides records of all PhD theses defended in French universities between 1988 and 2021. *Theses.fr* is a centralized public platform that systematically compiles data from university catalogs across France, sourced through library and documentation services within higher education and research institutions, establishing it as the most comprehensive and reliable platform for French PhD graduation.

The dataset is not immune to limitations. Data entry occurs manually at various stages, which introduces the potential for spelling inconsistencies. Furthermore, certain theses may go unreported due to a lack of submission by graduates, loss, or failure to meet quality control standards, which we estimate affects approximately 5% of theses each year. In addition, the processing of records is time-intensive, making the data for 2022 potentially incomplete. Additionally, an observed scarcity of records prior to 1988 suggests further underreporting. Consequently, we restrict our sample to the period from 1988 to 2021.

From an initial sample of 407,260 theses recorded between 1988 and 2021, we impose a series of exclusions to ensure data reliability. Theses supervised by more than two advisors—constituting roughly 2% of the dataset—are excluded, yielding a refined dataset of 399,118 observations. Additional filters are applied to exclude records with incomplete names for PhD candidates or supervisors, as well as cases with missing discipline information, resulting in a final dataset of 397,536 theses. At this stage, we exclude theses in medicine due to reliability concerns, which we discuss in detail in Section C.2, leaving a total of 340,073 observations.

For each thesis, we gathered information on the research discipline, defense year, university affiliation, and full names of the PhD student and supervisor(s). In the sections that follow, we detail the data-cleaning procedures applied to discipline and university affiliation, explain the exclusion of health and medical sciences, and outline our methodology for associating gender with first names.

C.1 Gender association

In this study, we determine the gender of both PhD students and supervisors based on first names. Our primary source is the INSEE database, which compiles first names assigned in France from 1900 to 2020, including the gender distribution for each name over the period 1940–2020. We focus on this range, assuming that the majority of PhD students in our dataset were born after 1940. For names associated with both genders, we establish a reliable gender ratio and retain only those names where one gender represents at least 95% of total occurrences; names below this threshold are treated as indeterminate. This process allows us to identify the gender for 305,187 out of 340,073 PhD student first names. Recognizing the limitations posed by foreign names, we supplement INSEE data with governmental databases from Australia, Canada, Spain, Sweden, the UK, and the US.

Through additional data collection from these international sources, we resolve the gender of an additional 9,246 PhD students. We further employ the methodology of Benveniste (2023), which classifies names based on the last two letters and the associated gender probability, allowing us to identify the gender of 3,004 more PhD students. In total, we successfully identify the gender of 317,437 doctoral students, covering 93% of the sample. Of the remaining 7%, 3% (8,166 names) represent names used by both genders without a clear distributional majority (e.g., Camille, Claude). Using the same approach, we successfully associate a gender for 95% of PhD supervisors.

C.2 Disciplines

The categorization of discipline fields in *Thèses.fr* is imprecise, partly due to manual data entry. The database originally contained around 22,000 unique entries for the discipline variable, which we grouped into twenty-two subcategories and further into four broader categories based on the Australian and New Zealand Standard Research Classification (ANZSRC). To classify these entries, we adopted a keyword-based approach, manually associating each entry with relevant discipline categories. We began by filtering with specific keywords unique to each category, as illustrated in the following examples:

Example
"CHIMIE ORGANIQUE" for "Chemical Sciences"
"INFORMATIQUE" for "Information, computing and Communication Sciences"
"SCIENCES BIOLOGIQUES" for "Biological Sciences" ...

Following this, we applied progressively broader keywords, carefully verifying that each association was accurate to avoid misclassification. For example, general keywords like "MAGNETISME," "LANGUES," and "VEGETAL" were used, corresponding to "Physical Sciences," "LanguAgeand Culture," and "Biological Sciences," respectively.

In cases of ambiguous or unknown disciplines, we examined thesis titles and applied the same keyword methodology. Despite these efforts, discipline association may still contain errors, especially for multidisciplinary theses that we must assign to a single category. To account for this, we created four overarching categories to group similar subjects: Humanities and Law, Biological and Earth Sciences, Sciences, Technology and Engineering, and Social Sciences.

Drop *Health and Medical Sciences* discipline. In this section, we discuss the unreliability of Health and Medical Sciences thesis data prior to the 2000s. Our analysis identified notable irregularities in medical theses data, particularly around 1994. We traced the origin of these discrepancies to the data selection mechanism in *Thèses.fr*, which automatically selects defended doctoral theses and excludes documents not categorized as such. However, in the French health sciences domain, "*thèses d'exercice*" - theses defended to obtain a State Diploma of Doctor required for medical practice—are often included. These are distinct from doctoral theses intended to confer the national diploma of doctor (*diplôme national de doctorat*). Unfortunately, during data import into *Thèses.fr*, a substantial number of *thèses d'exercice* were incorrectly labeled as doctoral theses, introducing bias.

Figure C3 displays the number of theses defended in health and medical sciences since 1988, showing that institutions began systematically distinguishing between doctoral theses and *thèses d'exercice* around the early 2000s. As we aim to focus on theses from before 2000, we must exclude medical theses from our sample to avoid biasing our study.

C.3 University

In recent years, French universities have been undergoing a series of institutional mergers, intended to enhance their international visibility and competitiveness[13]. To ensure consistency in our analysis, we standardized university codes following the documentation

[13]https://www.enseignementsup-recherche.gouv.fr/fr/premier-bilan-des-fusions-d-universites-realisees-entre-2009-et-2017-47515

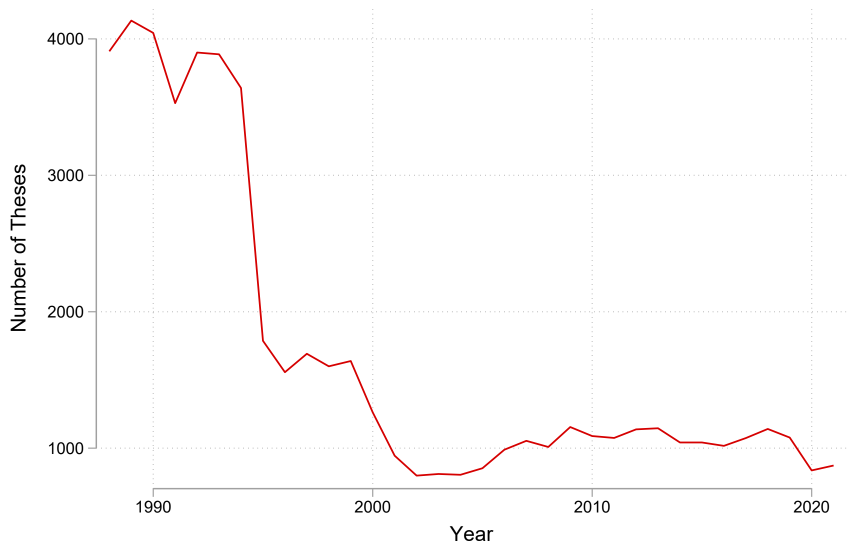


Figure C3: Number of Thesis by Year of Defense in Health and Medical Sciences

provided by Thèses.fr[14] and tracked changes in institutional names over time. Between 2007 and 2020, 26 new universities were established through the consolidation of 76 existing institutions. For example, in 2013, Aix-Marseille University was formed by merging Aix-Marseille 1, Aix-Marseille 2, and Aix-Marseille 3.

In certain cases, however, institutions have subsequently split, complicating the distinction between former codes. In such instances, it is more practical to apply a single code for universities that have separated, even at the cost of some specificity. For example, the University of Paris-Saclay was initially formed in 2015 as a merger of 11 institutions, only to divide into two distinct entities by the end of 2019.

Table C7, C8, and C9 provide a detailed list of all universities and their coding changes, while Table C10 covers the National Institutes of Polytechnics, and Table C11 presents the Higher Education Establishments. Each institution is listed with its associated code and any historical coding changes from 1988 to 2021. Any changes or codes appearing before or after this period are not documented. A blank description indicates no changes during the specified timeframe.

[14] https://documentation.abes.fr/guide/html/regles/CodesUnivEtab.htm

Code	University	Description
AGUY+ANTI+YANE*	Antilles-Guyane	ANTI and YANE since 2015
AIX1	Aix-Marseille 1	See AIXM since 2012
AIX2	Aix-Marseille 2	See AIXM since 2012
AIX3	Aix-Marseille 3	See AIXM since 2012
AIXM	Aix-Marseille	Creation 2012
AMIE	Amiens	
ANGE	Angers	
ANTI	Antilles	Creation 2015
ARTO	Artois	
AVIG	Avignon	
AZUR (=COAZ)**	Univ. Côte d'Azur (ComUE)	Creation 2016 then, changing code in 2020
BELF	Belfort Montbéliard	See UBFC since 2017
BESA	Besançon	See UBFC since 2017
BOR1 + BOR4***	Bordeaux 1 + 4	See BORD since 2014
BOR2	Bordeaux 2	See BORD since 2014
BOR3	Bordeaux 3	See BORD since 2014
BORD	Bordeaux	Creation 2014
BRES	Brest - Bretagne occidentale	
CAEN	Caen	See NORM since 2017
CERG (=CYUN)	Cergy-Pontoise	Changing code CYUN in 2020
CHAM	Chambéry	See GREN since 2010
CLF1	Clermont-Ferrand 1	See CLFA since 2021
CLF2	Clermont-Ferrand 2	See CLFA since 2021
CLFA (=UCFA)	Univ. Clermont Auvergne	Changing code UCFA in 2020
COMP	Compiègne	
CORT	Corte	
DIJO	Dijon	See UBFC since 2017
DUNK	Littoral Dunkerque	
EVRY	Evry Val d'Essonne	See SACL since 2015
GRAL	Univ. Grenoble Alpes	
GRE1	Grenoble 1	See GREN since 2010
GRE2	Grenoble 2	See GREN since 2010
GRE3	Grenoble 3	See GREN since 2010
GREN(=GREA = GRAL)	Grenoble	Changing code in 2015, 2020
LARE	La Réunion	
LARO	La Rochelle	
LEHA	Le Havre	See NORM since 2017
LEMA	Le Mans	

Table C7: Universities

All the code of the universities associated with their name and the evolution of their code over the years. We focus on the period 1988 to 2021, any changes and code that appears before or after are taken into account. If the description is empty, it means that there is no change during the period. * Guyane and Antilles were part of the same university at the beginning and then split, so we have to do only one university with all(because we don't know who was in which university); ** The sign equal, when the code name changed but represents the same university; *** BOR4 since 1995 for Law, Social Sciences and politics, Economics and Management theses), so we have to merge the two universities

Code	University	Description
LIL1	Lille 1	See LILU since 2018
LIL2	Lille 2	See LILU since 2018
LIL3	Lille 3	See LILU since 2018
LILU	Univ.polfLille	Creation 2018
LIMO	Limoges	
LORI	Lorient-Bretagne sud	
LORR	Univ. de Lorraine	Creation 2012
LYO1	Lyon 1	See LYSE since 2015
LYO2	Lyon 2	See LYSE since 2015
LYO3	Lyon 3	See LYSE since 2015
LYSE	Lyon (COMUE)	Creation 2015
MARN	Marne la Vallée	See PEST since 2008
METZ	Metz	See LORR since 2012
MON1	Montpellier 1	See MONT since 2015
MON2	Montpellier 2	See MONT since 2015
MON3	Montpellier 3	
MONT	Montpellier	Creation 2015
MULH	Mulhouse	
NAN1	Nancy 1	See LORR since 2012
NAN2	Nancy 2	See LORR since 2012
NANT	Nantes	
NCAL	Nouvelle Calédonie	
NICE	Nice	See AZUR since 2016
NIME	Nîmes	
NORM	Normandie (COMUE)	Creation 2017
PA01	Paris 1	
PA02	Paris 2	
PA03	Paris 3	See USPC between 2015-2019
PA04	Paris 4	See SORU since 2018
PA05	Paris 5	See USPC between 2015-2019
		See UNIP since 2019
PA06	Paris 6	See SORU since 2018
PA07	Paris 7	See USPC between 2015-2019
		See UNIP since 2019
PA08	Paris 8	
PA09	Paris 9	See PSLE since 2016
PA10	Paris 10	
PA11	Paris 11	See SACL since 2015
PA12	Paris 12	See PEST between 2008-2020
PA13	Paris 13	See USPC between 2015-2019

Table C8: Universities
All the code of the universities associated with their name and the evolution of their code over the years. We focus on the period 1988 to 2021, any changes and code that appears before or after are taken into account. If the description is empty, it means that there is no change during the period. * Nouvelle Calédonie and Polynésie française were part of the same university at the beginning and then split, so we have to use only one code with both as we can't distinguish them. ** PEST changed its name in 2015 to PESC

Code	University	Description
PACI +NCAL+POLF*	Pacifique	See NCAL and POLF since 1999
PAUU	Pau	
PERP	Perpignan	
PEST(=PESC)**	Paris Est (COMUE)	
POIT	Poitiers	
POLF	Polynésie française	
REIM	Reims	
REN1	Rennes 1	
REN2	Rennes 2	
ROUE	Rouen	
SACL+UPAS +IPPA+IAVF*	Univ. Paris-Saclay (ComUE)	Creation in 2015
SORU	Sorbonne Univ.	
STET	Saint-Etienne	See LYSE since 2015
STR1	Strasbourg 1	See STRA since 2009
STR2	Strasbourg 2	See STRA since 2009
STR3	Strasbourg 3	See STRA since 2009
STRA	Strasbourg	Creation 2009
TOU1	Toulouse 1	
TOU2	Toulouse 2	
TOU3	Toulouse 3-Ec. nationale vétérinaire	
TOUL	Toulon	
TOUR	Tours	
TROY	Troyes	
UBFC	Bourgogne Franche-Comté	Creation 2017
UCFA	Univ. Clermont-Auvergne	
UEFL	Univ. Gustave Eiffel	
UNIP	Univ. de Paris	Creation 2019
UPHF	Univ. Polytech. Hauts-de-France - Valenciennes	
USPC+PA03+PA13 +INAL+UNIP**	Sorbonne Paris Cité	Creation in 2019
VALE	Valenciennes	See UPHF since 2019
VERS	Versailles St Quentin en Yvelines	See SACL since 2015
YANE	Guyane	Creation 2015

Table C9: Universities

All the code of the universities associated with their name and the evolution of their code over the years. We focus on the period 1988 to 2021, any changes and code that appears before or after are taken into account. If the description is empty, it means that there is no change during the period. * IAVF is a new branch in 2016 and SACL was divided into UPAS and IPPA in 2019, as we can't distinguish, we use the same code for the three. ** There is a merge and then a split of universities, so we use one code for PA03, PA13, INAL, and UNIP only after 2019.

Code	Institute	Description
INPG	Institut national polytechnique - Grenoble	See GREN since 2009
INPL	Institut national polytechnique - Lorraine	
INPT	Institut national polytechnique - Toulouse	
IPPA	Institut Polytechnique de Paris	

Table C10: National Institute of Polytechnics
All the code of the universities associated with their name and the evolution of their code over the years. We focus on the period 1988 to 2021, any changes and code that appears before or after are taken into account. If the description is empty, it means that there is no change during the period.

Code	Establishment	Description
AGPT+EIAA +ENGR+INAP*	AgroParisTech	See SACL since 2015
CLIL	Centrale Lille Institut	
CNAM	Conservatoire national des arts et métiers	
CSUP	CentraleSupélec	See SACL since 2015
DENS	Ec. normale supérieure - Cachan	See SACL since 2015
ECAP	Ec. centrale des arts et manufactures de Paris	See SACL since 2015
ECDL	Ec. centrale de Lyon	See LYSE since 2015
ECDM	Ec. centrale de Marseille	
ECDN	Ec. centrale de Nantes	See CLIL since 2020
ECLI	Ec. centrale de Lille	See CLIL since 2020
EHEC	Ec. des hautes études commerciales	See SACL since 2015
EHES	Ec. des hautes études en sciences sociales	
EIAA	Ec. nationale supérieure des industries alimentaires - Massy	See AGPT since 2007-
EMAC	Ec. nationale des Mines d'Albi-Carmaux	
EMAL	IMT Mines Alès	
EMNA	Ec. des Mines de Nantes	See IMTA since 2017
EMSE	Ec. nationale supérieure des Mines - Saint-Etienne	
ENAM	Ec. nationale supérieure d'arts et métiers	See HESA since 2020
ENCM	Ec. nationale supérieure de chimie de Montpellier	
ENCP	Ec. nationale des chartes	
ENCR	Ec. nationale supérieure de chimie de Rennes	
ENGR	Ec. nationale du génie rural, des eaux et forêts	See AGPT since 2007
ENIB	Ec. nationale d'ingénieurs de Brest	
ENIS	Ec. nationale d'ingénieurs de Saint-Etienne	See LYSE since 2015
ENMP	Ec. nationale supérieure des Mines - Paris	See PSLE since 2016
ENPC	Ec. nationale des ponts et chaussées	See PEST since 2008
ENSL	Ec. normale supérieure (sciences) - Lyon	See LYSE since 2015
ENSR	Ec. normale supérieure de Rennes	
ENST	Ec. nationale supérieure des télécommunications	See SACL since 2015
ENSU	Ec. normale supérieure- Paris (rue d'Ulm)	See PSLE since 2016
ENTA	Ec. nationale supérieure de techniques avancées Bretagne	
ENTP	Ec. nationale des travaux publics	See LYSE since 2015
EPHE	Ec. pratique des hautes études	See PSLE since 2016
EPXX	Ec. polytechnique	See SACL since 2015
ESAE	ISAE	
ESEC	Ec. supérieure des sciences économiques et commerciales	
ESMA	Ec. nationale supérieure de mécanique et d'aérotechnique	
ESTA	Ec. nationale supérieure de techniques avancées	See SACL since 2015
GLOB	Institut de physique du Globe	See USPC since 2015
HESA	HESAM	
IAVF	Institut agronomique, vétérinaire et forestier de France - Paris	
IEPP	Institut d'études politiques - Paris	
IMTA	Ec. nationale supérieure Mines-Télécom Atlantique Bretagne Pays de la Loire	
INAL	Institut national des langues et civilisations orientales (INALCO)	See USPC since 2015
INAP	Institut national d'agronomie - Paris Grignon	See AGPT since 2007
IOTA	Institut d'optique théorique et appliquée - Palaiseau	SACL UPAS
ISAB	Institut national des sciences appliquées Val de Loire - Bourges	
ISAL	Institut national des sciences appliquées - Lyon	See LYSE since 2015
ISAM	Institut national des sciences appliquées - Rouen	See NORM since 2017
ISAR	Institut national des sciences appliquées - Rennes	
ISAT	Institut national des sciences appliquées - Toulouse	
MNHN	Museum d'histoire naturelle	
MTLD	Ec. nationale supérieure Mines-Télécom Lille Douai	
NSAI	Ec. nationale de la Statistique et de l'Analyse de l'Information - Rennes	
NSAM	SupAgro - Montpellier	
NSAR	Agrocampus Ouest - Rennes	
OBSP	Observatoire de Paris	See PSLE since 2016
ONIR	Ec. nationale vétérinaire - Nantes	
ORLE		
PSLE(=UPSL)	Paris Sciences et Lettres (ComUE)	Creation 2016
TELB	Ec. nationale supérieure des TelecompolfBretagne - Brest	See IMTA since 2017
TELE	Institut national des télécommunications	See SACL since 2015

Table C11: Higher Education Establishment
All the code of the universities associated with their name and the evolution of their code over the years. We focus on the period 1988 to 2021, any changes and code that appears before or after are taken into account. If the description is empty, it means that there is no change during the period. * EIAA+ENGR+INAP merged to become AGPT in 2007 we use one code for the three. ** Change code in 2020

D Robustness

Table D12: Determinants of Application Behavior: Candidate–Job Dyads - Distance in km

	(1)	(2)	(3)
Dependent variable:		*Apply to position*	
Female	0.00197	-0.00277	-
	(0.00183)	(0.00203)	-
Distance (km)	-0.0000748***	-	-0.0000744***
	(0.00000199)	-	(0.00000186)
Distance (km) × Female	-0.0000206***	-0.00000749**	-0.0000216***
	(0.00000294)	(0.00000379)	(0.00000259)
Adj R^2	0.19	0.20	0.30
Controls	yes	yes	yes
Fixed effects	$U_i \times t \times f + U_j \times t \times f$	$U_i \times U_j \times t \times f$	$i \times (t \times f) + j \times (t \times f)$
Observations	2,287,422	2,162,136	2,286,953

Notes: *Apply* is a binary variable equal to 1 if the candidate applied to a specific job. Each observation represents a dyad between a candidate and a potential job opening. *Distance* is the geographical distance between the job and the candidate's PhD institution. All regressions include controls for age, publication metrics, and supervisor gender. Fixed effects vary across specifications and are indicated in the "Fixed effects" row: U_i denotes the university of candidate i, U_j the university of the job j, t the year, and f the field. i and j denote candidate and job identifiers, respectively. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D13: Determinants of Application Behavior: Candidate–Job Dyads - Commuting Time

	(1)	(2)	(3)
Dependent variable:		*Apply to position*	
Female	0.00123	-0.00292	-
	(0.00178)	(0.00195)	-
Commuting time (min)	-0.000127***	-	-0.000124***
	(0.00000364)	-	(0.00000333)
Commuting time (min) × Female	-0.0000346***	-0.0000132**	-0.0000429***
	(0.00000517)	(0.00000662)	(0.00000413)
Adj R^2	0.19	0.20	0.30
Controls	yes	yes	yes
Fixed effects	$U_i \times t \times f + U_j \times t \times f$	$U_i \times U_j \times t \times f$	$i \times (t \times f) + j \times (t \times f)$
Observations	2,287,422	2,162,136	2,286,968

Notes: *Apply* is a binary variable equal to 1 if the candidate applied to a specific job. Each observation represents a dyad between a candidate and a potential job opening. *Commuting time* combines train travel time (from official SNCF timetables), road travel time (based on routing algorithms), and AI-based predictions for less connected pairs. Details of the construction are provided in Appendix Section D.2@. between the job and the candidate's PhD institution. All regressions include controls for age, publication metrics, and supervisor gender. Fixed effects vary across specifications and are indicated in the "Fixed effects" row: U_i denotes the university of candidate i, U_j the university of the job j, t the year, and f the field. i and j denote candidate and job identifiers, respectively. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D14: Determinants of Application Behavior: Candidate–Job Dyads - controls: age at PhD and time since graduation

	(1)	(2)
Dependent variable:	*Apply to position*	
Female	0.00710**	0.000582
	(0.00279)	(0.00323)
ln(Distance)	-0.0127***	-
	(0.000319)	-
ln(Distance) $\times$ Female	-0.00236***	-0.00113**
	(0.000439)	(0.000533)
Adj R^2	0.19	0.20
Controls	yes	yes
Fixed effects	$U_i \times t \times f + U_j \times t \times f$	$U_i \times U_j \times t \times f$
Observations	2,287,422	2,162,136

Notes: *Apply* is a binary variable equal to 1 if the candidate applied to a specific job. Each observation represents a dyad between a candidate and a potential job opening. *ln(Distance)* is the logarithm of the geographical distance between the job and the candidate's PhD institution. All regressions include controls for age, publication metrics, and supervisor gender. Fixed effects vary across specifications and are indicated in the "Fixed effects" row: U_i denotes the university of candidate i, U_j the university of the job j, t the year, and f the field. i and j denote candidate and job identifiers, respectively. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D15: Determinants of Application Behavior: Candidate–Job Dyads - Excluding Paris candidates

	(1)	(2)	(3)
Dependent variable:		*Apply to position*	
Female	0.0208***	0.000107	-
	(0.00436)	(0.00510)	-
ln(Distance)	-0.0154***	-	-0.0153***
	(0.000409)	-	(0.000397)
ln(Distance) $\times$ Female	-0.00482***	-0.00118	-0.00513***
	(0.000686)	(0.000837)	(0.000651)
Adj R^2	0.20	0.21	0.30
Controls	yes	yes	yes
Fixed effects	$U_i \times t \times f + U_j \times t \times f$	$U_i \times U_j \times t \times f$	$i \times (t \times f) + j \times (t \times f)$
Observations	1,576,906	1,470,253	1,576,684

Notes: *Apply* is a binary variable equal to 1 if the candidate applied to a specific job. Each observation represents a dyad between a candidate and a potential job opening. *ln(Distance)* is the logarithm of the geographical distance between the job and the candidate's PhD institution. All regressions include controls for age, publication metrics, and supervisor gender. Fixed effects vary across specifications and are indicated in the "Fixed effects" row: U_i denotes the university of candidate i, U_j the university of the job j, t the year, and f the field. i and j denote candidate and job identifiers, respectively. Qualified candidates from Paris are excluded. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D16: Determinants of Application Behavior: Candidate–Job Dyads - Subsamples Based on Application Timing

	(1)	(2)	(3)
		Apply to position	
Panel A: Candidates who applied at least once in their career			
Female	0.00771**	0.00174	-
	(0.00334)	(0.00390)	-
ln(Distance)	-0.0171***	-	-0.0172***
	(0.00040)	-	(0.00039)
ln(Distance) × Female	-0.00255***	-0.00139**	-0.00249***
	(0.00054)	(0.00066)	(0.00050)
Adj R^2	0.20	0.20	0.30
Observations	1,726,884	1,601,520	1,726,649
Panel B: Candidates who applied in year of first qualification			
Female	0.00692*	0.00111	-
	(0.00371)	(0.00438)	-
ln(Distance)	-0.0214***	-	-0.0214***
	(0.00048)	-	(0.00047)
ln(Distance) × Female	-0.00246***	-0.00132*	-0.00263***
	(0.00062)	(0.00078)	(0.00060)
Adj R^2	0.20	0.21	0.28
Observations	1,390,720	1,272,226	1,390,599
Panel C: Second year after qualification			
Female	0.00391	0.00196	-
	(0.00252)	(0.00294)	-
ln(Distance)	-0.00677***	-	-0.00687***
	(0.00023)	-	(0.00021)
ln(Distance) × Female	-0.00157***	-0.00118***	-0.00133***
	(0.00037)	(0.00045)	(0.00030)
Adj R^2	0.147	0.132	0.297
Observations	2,021,493	1,907,890	2,020,987
Panel D: Third year after qualification			
Female	0.00109	-0.00035	-
	(0.00209)	(0.00249)	-
ln(Distance)	-0.00435***	-	-0.00432***
	(0.00019)	-	(0.00018)
ln(Distance) × Female	-0.00091***	-0.00062	-0.00098***
	(0.00031)	(0.00040)	(0.00025)
Adj R^2	0.120	0.102	0.273
Observations	1,764,522	1,662,447	1,764,043
Panel E: Fourth year after qualification			
Female	-0.00006	-0.00306	-
	(0.00165)	(0.00194)	-
ln(Distance)	-0.00289***	-	-0.00281***
	(0.00016)	-	(0.00015)
ln(Distance) × Female	-0.00043*	0.00015	-0.00061***
	(0.00024)	(0.00030)	(0.00021)
Adj R^2	0.099	0.076	0.248
Observations	1,524,331	1,434,473	1,523,912
Controls	yes	yes	yes
Fixed effects	$U_i \times t \times f + U_j \times t \times f$	$U_i \times U_j \times t \times f$	$i \times (t \times f) + j \times (t \times f)$

Notes: *Apply* is a binary variable equal to 1 if the candidate applied to a specific job. Each observation represents a dyad between a candidate and a potential job opening. *ln(Distance)* is the logarithm of the geographical distance between the job and the candidate's PhD institution. All regressions include controls for age, publication metrics, and supervisor gender. Fixed effects vary across specifications and are indicated above. Standard errors are clustered by Discipline × Candidate Univ × Year. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D17: Determinants of Application Behavior: Candidate–Job Dyads - Heterogeneity by Field

	(1)	(2)	(3)
		Apply to position	
Panel A: Qualified in biological and earth sciences			
Female	0.00039	0.00079	-
	(0.00401)	(0.00464)	-
ln(Distance)	-0.00790***	-	-0.00788***
	(0.00055)	-	(0.00057)
ln(Distance) $\times$ Female	-0.00058	-0.00062	-0.00062
	(0.00066)	(0.00077)	(0.00068)
Adj R^2	0.08	0.09	0.13
Observations	245,228	220,687	244,527
Panel B: Qualified in humanities			
Female	-0.00468	-0.00687	-
	(0.00385)	(0.00429)	-
ln(Distance)	-0.0121***	-	-0.0122***
	(0.00056)	-	(0.00054)
ln(Distance) $\times$ Female	-0.00071	-0.00020	-0.00054
	(0.00062)	(0.00072)	(0.00060)
Adj R^2	0.09	0.09	0.18
Observations	611,387	569,559	609,639
Panel C: Qualified in STEM			
Female	0.00999**	0.00701	-
	(0.00465)	(0.00543)	-
ln(Distance)	-0.0100***	-	-0.00998***
	(0.00036)	-	(0.00035)
ln(Distance) $\times$ Female	-0.00211***	-0.00157*	-0.00223***
	(0.00074)	(0.00089)	(0.00064)
Adj R^2	0.10	0.10	0.21
Observations	1,045,478	987,343	1,044,110
Panel D: Qualified in social sciences			
Female	0.0128	0.00974	-
	(0.0103)	(0.0119)	-
ln(Distance)	-0.0255***	-	-0.0263***
	(0.00126)	-	(0.00116)
ln(Distance) $\times$ Female	-0.00454***	-0.00409**	-0.00289**
	(0.00160)	(0.00195)	(0.00136)
Adj R^2	0.27	0.28	0.41
Observations	382,761	362,659	381,733
Controls	yes	yes	yes
Fixed effects	$U_i \times t \times f + U_j \times t \times f$	$U_i \times U_j \times t \times f$	$i \times (t \times f) + j \times (t \times f)$

Notes: *Apply* is a binary variable equal to 1 if the candidate applied to a specific job. Each observation represents a dyad between a candidate and a potential job opening. *ln(Distance)* is the logarithm of the geographical distance between the job and the candidate's PhD institution. All regressions include controls for age, publication metrics, and supervisor gender. Fixed effects vary across specifications and are indicated in the original field-specific tables. Standard errors clustered by Discipline $\times$ Candidate Univ $\times$ Year. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D18: Determinants of Application Behavior: Candidate–Job Dyads - Heterogeneity Analysis

	(1)	(2)
Dependent variable:	*Apply to position*	
Age	$\geq$ median	$<$ median
ln(Distance)	-0.0121***	-0.0131***
	(0.000433)	(0.000379)
ln(Distance) × Female	-0.00303***	-0.00186***
	(0.000645)	(0.000494)
Years since PhD	$\geq$ median	$<$ median
ln(Distance)	-0.00988***	-0.0149***
	(0.000348)	(0.000449)
ln(Distance) × Female	-0.00150***	-0.00288***
	(0.000496)	(0.000585)
Total AIS	$= 0$	> 0
ln(Distance)	-0.0155***	-0.0100***
	(0.000474)	(0.000347)
ln(Distance) × Female	-0.00197***	-0.00149***
	(0.000563)	(0.000519)
Number Publications	$= 0$	> 0
ln(Distance)	-0.0170***	-0.0104***
	(0.000589)	(0.000310)
ln(Distance) × Female	-0.00117*	-0.00207***
	(0.000675)	(0.000469)
Controls, FE	yes	yes

Notes: *Apply* is a binary variable equal to 1 if the candidate applied to a specific job. Each observation represents a dyad between a candidate and a potential job opening. *ln(Distance)* is the logarithm of the geographical distance between the job and the candidate's PhD institution. All regressions include controls for age, publication metrics, and supervisor gender. Fixed effects vary across Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.1 Individual-level Application Behavior

Table D19: Gender Differences in Application Patterns by Distance to Job Offers - Pseudo-Maximum Likelihood

	Applications to Nearby Jobs ($\leq$ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*Near Apps*	*Near Apps*	*Near Apps*	*Near Apps*	*Near Apps*	*Near Apps*
Female	-0.0528**	-0.0639***	-0.0532**	-0.0712**	-0.0917***	-0.102***
	(0.0241)	(0.0240)	(0.0259)	(0.0285)	(0.0284)	(0.0283)
Near offers	0.0939***	0.0942***	0.0755***			
	(0.00206)	(0.00206)	(0.00444)			
Female $\times$ Near offers	0.00506***	0.00531***	0.00351*			
	(0.00184)	(0.00183)	(0.00205)			
Far offers				0.0297***	0.0291***	0.0163***
				(0.00131)	(0.00130)	(0.00316)
Female $\times$ Far offers				-0.0000862	0.0000112	0.000114
				(0.000672)	(0.000663)	(0.000649)
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	66628.00	66628.00	57979.00	67824.00	67824.00	64992.00

Notes: The dependent variable is the number of applications, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. Estimated using Poisson Pseudo-Maximum Likelihood (PPML) regression. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D20: Gender Differences in Application Patterns by Distance to Job Offers - Controls

	Applications to Nearby Jobs (≤ 100km)		*Applications to Distant Jobs (>100km)*	
	(1)	(2)	(3)	(4)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*
Female	-0.0228***	-0.00908**	-0.0341***	-0.0319***
	(0.00404)	(0.00412)	(0.00867)	(0.00896)
Near offers	0.0280***	0.0299***		
	(0.000789)	(0.00124)		
Female × Near offers	0.00455***	0.00131		
	(0.000964)	(0.00101)		
Far offers			0.0172***	0.0107***
			(0.000730)	(0.00154)
Female × Far offers			-0.0000998	-0.000258
			(0.000348)	(0.000356)
Adj R-squared	0.33	0.38	0.34	0.35
Controls	yes	yes	yes	yes
Fields X Year FE	yes	yes	yes	yes
Fields X Univ PhD FE		yes		yes
Observations	68258	67617.00	68258.00	67617.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age at PhD defense, time since PhD graduation, publication metrics, supervisor gender, and supervisor productivity. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D21: Gender Differences in Application Patterns by Distance to Job Offers - Excluding Paris

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.0183***	-0.0194***	-0.00876**	-0.0368***	-0.0454***	-0.0441***
	(0.00414)	(0.00418)	(0.00428)	(0.0112)	(0.0112)	(0.0117)
Near offers	0.0270***	0.0270***	0.0397***			
	(0.00118)	(0.00118)	(0.00184)			
Female × Near offers	0.00631***	0.00635***	0.00152			
	(0.00176)	(0.00176)	(0.00179)			
Far offers				0.0159***	0.0160***	0.00738***
				(0.00124)	(0.00124)	(0.00242)
Female × Far offers				0.0000701	0.000107	-0.000107
				(0.000419)	(0.000416)	(0.000427)
Adj R-squared	0.26	0.25	0.30	0.32	0.33	0.35
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	45385.00	45385.00	44882.00	45385.00	45385.00	44882.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. This sample exclude candidates from Paris. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D22: Gender Differences in Application Patterns by Distance to Job Offers -Apply at least once in the career

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.0164***	-0.0198***	-0.00613	-0.0183*	-0.0283***	-0.0237**
	(0.00495)	(0.00497)	(0.00512)	(0.0107)	(0.0107)	(0.0113)
Near offers	0.0392***	0.0393***	0.0384***			
	(0.000892)	(0.000894)	(0.00145)			
Female × Near offers	0.00421***	0.00431***	0.000976			
	(0.00105)	(0.00105)	(0.00109)			
Far offers				0.0198***	0.0196***	0.0140***
				(0.000827)	(0.000824)	(0.00180)
Female × Far offers				-0.000471	-0.000503	-0.000730*
				(0.000390)	(0.000386)	(0.000402)
Adj R-squared	0.40	0.40	0.45	0.37	0.37	0.39
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	49648.00	49648.00	48959.00	49648.00	49648.00	48959.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. This sample is restricted to candidates who applied at least once for a position during their career. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D23: Gender Differences in Application Patterns by Distance to Job Offers -Apply at least once in first year qualification

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.0133**	-0.0151**	-0.00347	-0.0301**	-0.0372***	-0.0317**
	(0.00626)	(0.00628)	(0.00634)	(0.0120)	(0.0120)	(0.0130)
Near offers	0.0476***	0.0476***	0.0440***			
	(0.000991)	(0.000993)	(0.00159)			
Female × Near offers	0.00288**	0.00296**	0.000309			
	(0.00115)	(0.00115)	(0.00116)			
Far offers				0.0250***	0.0248***	0.0210***
				(0.000851)	(0.000849)	(0.00192)
Female × Far offers				-0.000479	-0.000475	-0.000708*
				(0.000389)	(0.000387)	(0.000411)
Adj R-squared	0.46	0.45	0.50	0.50	0.41	0.41
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	37755.00	37755.00	36987.00	37755.00	37755.00	36987.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. This sample is restricted to candidates who applied at least once during the first year of qualification. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D24: Gender Differences in Application Patterns by Distance to Job Offers - Biological and Earth Sciences

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.00437	-0.00297	-0.00501	-0.0258**	-0.0236**	-0.0180
	(0.00380)	(0.00380)	(0.00405)	(0.0115)	(0.0116)	(0.0114)
Near offers	0.0105***	0.0105***	0.0181***			
	(0.00147)	(0.00147)	(0.00229)			
Female × Near offers	0.00234	0.00224	0.00246			
	(0.00179)	(0.00178)	(0.00195)			
Far offers				0.00834***	0.00824***	0.00280
				(0.00130)	(0.00130)	(0.00224)
Female × Far offers				-0.000417	-0.000425	-0.000374
				(0.000712)	(0.000712)	(0.000706)
Adj R-squared	0.07	0.07	0.08	0.06	0.07	0.10
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	12531.00	12531.00	12200.00	12531.00	12531.00	12200.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. This sample is restricted to candidates in Biological and Earth Sciences. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D25: Gender Differences in Application Patterns by Distance to Job Offers - Humanities

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.00233	-0.000639	-0.00112	-0.0384***	-0.0326**	-0.0348**
	(0.00631)	(0.00631)	(0.00676)	(0.0139)	(0.0139)	(0.0148)
Near offers	0.0200***	0.0201***	0.0240***			
	(0.000903)	(0.000905)	(0.00168)			
Female × Near offers	0.00212	0.00206	0.00156			
	(0.00155)	(0.00155)	(0.00160)			
Far offers				0.0155***	0.0154***	0.00736***
				(0.000957)	(0.000955)	(0.00221)
Female × Far offers				0.000552	0.000507	0.000352
				(0.000462)	(0.000461)	(0.000476)
Adj R-squared	0.17	0.17	0.20	0.15	0.15	0.17
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	26485.00	26485.00	26075.00	26485.00	26485.00	26075.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. This sample is restricted to candidates in Humanities. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D26: Gender Differences in Application Patterns by Distance to Job Offers -STEM

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.0133**	-0.0151**	-0.00347	-0.0301**	-0.0372***	-0.0317**
	(0.00626)	(0.00628)	(0.00634)	(0.0120)	(0.0120)	(0.0130)
Near offers	0.0476***	0.0476***	0.0440***			
	(0.000991)	(0.000993)	(0.00159)			
Female × Near offers	0.00288**	0.00296**	0.000309			
	(0.00115)	(0.00115)	(0.00116)			
Far offers				0.0250***	0.0248***	0.0210***
				(0.000851)	(0.000849)	(0.00192)
Female × Far offers				-0.000479	-0.000475	-0.000708*
				(0.000389)	(0.000387)	(0.000411)
Adj R-squared	0.46	0.45	0.50	0.50	0.41	0.41
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	37755.00	37755.00	36987.00	37755.00	37755.00	36987.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. This sample is restricted to candidates in STEM. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D27: Gender Differences in Application Patterns by Distance to Job Offers - Social Sciences

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.00112	-0.00822	-0.0113	-0.0613	-0.0606	-0.0651
	(0.0237)	(0.0237)	(0.0232)	(0.0470)	(0.0464)	(0.0514)
Near offers	0.0478***	0.0484***	0.0409***			
	(0.00235)	(0.00234)	(0.00449)			
Female × Near offers	0.000796	0.00104	0.00105			
	(0.00295)	(0.00294)	(0.00285)			
Far offers				0.0290***	0.0278***	0.0141**
				(0.00219)	(0.00215)	(0.00569)
Female × Far offers				0.00000645	-0.000348	-0.000426
				(0.000940)	(0.000920)	(0.000982)
Adj R-squared	0.36	0.37	0.44	0.28	0.30	0.32
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	6477.00	6477.00	6262.00	6477.00	6477.00	6262.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. This sample is restricted to candidates in Social sciences. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D28: Gender Differences in Application Patterns by Distance to Job Offers - Same city

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.0121***	-0.0133***	-0.00472	-0.0301***	-0.0370***	-0.0352***
	(0.00338)	(0.00338)	(0.00334)	(0.00877)	(0.00877)	(0.00910)
Near offers	0.0357***	0.0358***	0.0352***			
	(0.00108)	(0.00108)	(0.00147)			
Female × Near offers	0.00417***	0.00420***	0.00140			
	(0.00120)	(0.00120)	(0.00122)			
Far offers				0.0179***	0.0176***	0.0117***
				(0.000886)	(0.000883)	(0.00183)
Female × Far offers				0.0000561	0.0000710	-0.0000874
				(0.000335)	(0.000333)	(0.000342)
Adj R-squared	0.35	0.35	0.40	0.33	0.33	0.35
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	68258.00	68258.00	67617.00	68258.00	68258.00	67617.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D29: Gender Differences in Application Patterns by Distance to Job Offers - Same region

	Applications to Nearby Jobs (≤ 100km)			*Applications to Distant Jobs (>100km)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(near apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*	*ln(far apps + 1)*
Female	-0.0234***	-0.0258***	-0.0146***	-0.0240***	-0.0308***	-0.0303***
	(0.00453)	(0.00454)	(0.00474)	(0.00859)	(0.00860)	(0.00889)
Near offers	0.0270***	0.0271***	0.0292***			
	(0.000820)	(0.000822)	(0.00123)			
Female × Near offers	0.00419***	0.00426***	0.00181*			
	(0.000994)	(0.000991)	(0.00104)			
Far offers				0.0174***	0.0172***	0.0114***
				(0.000807)	(0.000805)	(0.00153)
Female × Far offers				-0.000226	-0.000214	-0.000325
				(0.000353)	(0.000351)	(0.000359)
Adj R-squared	0.34	0.34	0.37	0.32	0.33	0.35
Controls		yes	yes		yes	yes
Fields X Year FE	yes	yes	yes	yes	yes	yes
Fields X Univ PhD FE			yes			yes
Observations	68258.00	68258.00	67617.00	68258.00	68258.00	67617.00

Notes: The dependent variable is the natural logarithm of the number of applications plus one, submitted by candidates, separately for nearby job offers (within 100km) and distant job offers (over 100km). Control variables include age, publication metrics, supervisor gender, and supervisor productivity. Standard errors are clustered by Discipline X Candidate Univ X Year. Significance levels are defined as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.2 Construction of the Commuting Time Variable

To complement great-circle distance as a measure of geographic frictions, I construct a variable for estimated commuting time between the PhD institution and the job location. This variable is designed to better capture realistic travel costs faced by candidates, accounting for transportation infrastructure and regional accessibility.

The commuting time is computed in several steps:

1. **Train-based commuting time.** I merge the dyadic dataset with an external dataset containing average train travel times between French cities, using information from the French national railway open data platform (data.sncf.com). The merge is based on year and city-to-city routes (e.g., "Lyon–Paris").

2. **Special adjustments for the Paris region.** For movements within the Île-de-France region, where suburban candidates frequently commute to central Paris, I assign a default value of 60 minutes, reflecting typical intra-regional commuting durations. This value is applied to both directions (from/to Paris). In a second step, I redefine all cities within Île-de-France as "Paris" to capture additional matches in the train time dataset. I re-merge the data and add 60 minutes to the retrieved travel time to account for average commuting from the broader metropolitan area.

3. **Fallback proxy using road travel time.** For remaining unmatched observations, I impute commuting time using a road-based proxy derived from great-circle distance. Assuming a speed of 90 km/h and inflating the straight-line distance by a factor of 1.2, I approximate round-trip travel time as follows:

$$\text{Commuting Time (min)} = 2 \times \left(\frac{1.2 \times \text{Distance (km)}}{90} \right) \times 60$$

   This provides a conservative estimate of round-trip driving time.