What Shaped Women's Invention during Industrialization?

Youssouf Merouani

2026-04-19

Abstract

The local economy predicted what nineteenth-century French women invented. Their life course predicted when. I link 1,076 women who patented between 1791 and 1900 to genealogical records to examine both questions. The linkage reveals a geographic distortion: 80 percent of patent addresses fall in Paris, but civil records show that many of these women lived and married elsewhere. Using birth departments rather than patent addresses, I find that women patented in industries where local women already worked: the gender composition of employment, not its total size, was the dominant predictor of sector choice. Married women first patented at a median age of 41.5, roughly 12 years later than single women. Among mothers, the first patent followed the last recorded birth by a median of 11 years, and no department-level measure of female economic activity detectably narrowed the marriage gap. Among mothers observable through age 50, women who came of age around the 1844 patent reform patented about four years earlier than earlier cohorts. Most of this shift reflects earlier completion of childbearing and a shorter post-childbearing gap consistent with the institutional changes that followed.

Keywords: Women inventors · Demography · Geography of innovation · Nineteenth-century France
JEL Codes: J16, N33, O31, J13, R12

1 Introduction

Between 1791 and 1900, 4,165 French women filed nearly 7,000 patents (Merouani, 2026; Merouani and Perrin, 2026b).[1] They invented in textiles, chemistry, clothing, machinery, and domestic goods. They patented as young as their twenties and as late as their seventies. Patent records identify what these women invented and when they filed, but they do not capture the local labor markets and family lives that may have shaped the direction and timing of invention. Why did a woman in Alsace patent a textile innovation while a woman in Provence patented a chemical process? What shaped the age at which they first filed?

I cannot answer these questions from patent files alone, because they do not contain the biographical evidence that Chanteux (2023, 2024) identify as missing from the study of French women's invention. I fill that biographical gap by linking patent records to genealogical sources, producing 1,076 demographic profiles that include birth years, marriage timing, geographic origins, children, and death years. The linkage also exposes a geographic distortion in the patent record: 80 percent of patent addresses fall in Paris, but civil records reveal that many of these women lived and married elsewhere in France. The gap arises because France's centralized filing system routed applications through Paris-based agents, whose addresses replaced inventors' own (Galvez-Behar, 2019). The result is a Paris bias in the administrative record, not a Parisian concentration of women's lives. I use birth locations rather than patent addresses to identify where inventors came from.

Gendered occupational knowledge, not industrial scale, predicted what women invented. If agglomeration alone mattered, women should have patented in the largest local sectors. If gendered occupational knowledge mattered, women should have patented in local sectors with a higher female employment share. I estimate a conditional logit comparing industries within each inventor's birth department. The data support the second prediction: the female employment share of a local sector predicted patenting in that sector, while total sector size did not. Marie Massart, born in Haut-Rhin where textiles dominated women's employment, patented a textile innovation. Marie

[1] This count includes only individual women and excludes female-led firms. When those firms are included, the total exceeds 4,300 entities.

Louise Trouilhas, born in Bouches-du-Rhône where the local soap industry put women in contact with industrial chemistry, patented a chemical process. What mattered was not the size of a nearby industry but the share of women in its workforce.

The local economy predicted what women invented, but marriage, childbearing, and cohort predicted when they patented. If local conditions had helped married women combine domestic duties with productive work, the marriage gap should have been smaller where female industrial employment was higher. It was not. Married women patented roughly 12 years later than single women, regardless of local female industrial employment. Nor did literacy, urbanization, wage gaps, or other department-level conditions predict earlier first patenting. But while the marriage gap held across departments, the age at first patent fell across cohorts. The generation that came of age around the 1844 patent reform patented about four years earlier than its predecessors. This shift coincided with the long-run French fertility transition (Perrin, 2022b). It also coincided with three institutional changes: annual installments for patent fees after 1844 (Galvez-Behar, 2019), expanding girls' education (Perrin, 2022a), and greater financial autonomy for married women after the 1880s (Lewis, 1980; McMillan, 2000).

This paper makes three contributions. First, I build a new dataset linking 1,076 nineteenth-century French women inventors to their genealogical records, to my knowledge the first such linkage in the French context. The data also show that patent addresses in centralized filing systems should not be treated as residential proxies, a finding relevant to any study that maps the geography of innovation from administrative records. Second, I use this dataset to extend the geography-of-innovation literature (Akcigit et al., 2017; Bell et al., 2019), which has focused on twentieth-century men, to women during industrialization. Third, I connect patent timing to the child-penalty literature (Kim and Moser, 2025; Kleven et al., 2019), documenting that among mothers, first patents typically followed the end of childbearing. The results also provide evidence on whether mothers withdrew from productive work or intensified it as children grew older (Tilly and Scott, 1989; You, 2020) and on how technological change restructured women's work (Humphries and Schneider, 2019; Schneider, 2026).

The geography and timing results are strong associations, not causal estimates. I measure local exposure by birth department, but the results are similar or stronger when I assign married inventors to their marriage department and unmarried inventors to their birth department, because marriage records usually place them closer to the patent date. The industrial structure comes from a single 1860 cross-section. The geography results therefore reflect which departments were industrialized at mid-century, not the specific industrial conditions each woman faced when she was young. The timing results describe women who eventually patented, not the full effect of marriage and childbearing on the probability of ever patenting.

Section 2 provides background on the French patent system and lays out the competing mechanisms. Section 3 describes the data and linking methodology. Section 4 establishes the geographic origins of women inventors. Section 5 examines whether local industry predicted the direction of women's patents. Section 6 documents life-cycle timing and its relationship to marriage and childbearing. Section 7 concludes.

2 Invention, work, and the life course

The French patent system was formally open to women, but filing a patent required fees, technical knowledge, and legal autonomy that women accumulated slowly and unevenly over their lives. Understanding who patented, in which sectors, and at what age requires separating two questions. The first is about direction: what shaped the sectors in which women patented? The second is about timing: what shaped when in their lives women filed their first patent?

2.1 The patent system

The revolutionary laws of 1791 created a registration regime that rejected preliminary examination of patent applications and imposed no formal restrictions by gender, age, or civil status. In practice, an unofficial examination by the Comité consultatif des Arts et des Manufactures operated before 1844, but the 1844 reform made the openness explicit: Article 5 declared the formalities applicable to "whoever wishes to take out a patent," and an accompanying instruction specified that the administration must issue patents even when the applicant was a married woman, a minor, or a legally incapacitated person (Merouani and Perrin, 2024). In Britain, by contrast, married women's property rights were restricted by coverture until the reforms of 1870 and 1882 (Khan, 2020).

The system was costly. Under the 1791 regime, a five-year patent cost 300 francs, roughly two hundred days' wages for an average mid-century worker (Galvez-Behar, 2019; Nuvolari et al., 2023). The 1844 reform raised total fees but spread payment over the patent's duration at 100 francs per year, a sum within reach of artisans and small entrepreneurs (Galvez-Behar, 2019). The 1844 Act also required patentees to exploit their invention in France; a patent that was not put into production could be revoked (Galvez-Behar, 2019). Filing was therefore not merely an intellectual claim but a commitment to commercialize. Merouani and Perrin (2026b) show that the 1844 reform is associated with a 40 percent increase in the likelihood of a patent being filed by a woman. Among women patentees with known occupations, 44 percent belonged to higher social classes, but 31 percent held medium-skill jobs and 25 percent came from lower-skilled backgrounds, a breadth that contrasts with late-nineteenth-century Sweden, where patenting was concentrated in a small industrial elite (Berger and Prawitz, 2024).

Applications were processed in Paris, often through professional agents who handled the filing procedures on behalf of inventors and whose addresses ended up on the patent record (Galvez-Behar, 2019). Patentees outside Paris routinely listed their Parisian agent's address rather than their own. Galvez-Behar (2019) reports that over 53 percent of patentees in the early 1840s gave Paris addresses; whether the same distortion applied to women inventors throughout the century has not been established.

The stakes of this distortion extend beyond the present sample because patent records are the standard source for mapping the geography of innovation (Akcigit et al., 2017; Nuvolari and Vasta, 2017). In countries with decentralized filing, such as Italy, patent addresses reliably indicate residence (Nuvolari and Vasta, 2017). In France, all filing passed through Paris, and Nuvolari et al. (2023), studying patents from 1791 to 1844, note that the resulting concentration in the Seine department likely overstates the capital's true share. Once Paris is set aside, the spatial distribution of patenting follows the North-East/South-West development gradient (Braudel, 1984; Dupin, 1826). Whether women's patents suffered the same geographic distortion across the full nineteenth century has not been examined.

Women's legal position under the Napoleonic Code was one of subordination to fathers and husbands, but practice diverged from theory. Married women needed their husbands' permission to manage property or conduct business, yet they developed strategies to work around these restrictions. Khan (2016), drawing on patent and exhibition records, shows that middle-class French women were extensively engaged in entrepreneurship and innovation, and that their commercial efforts were significantly enhanced by association with family firms. This pattern was not confined to patenting: studies of Lille and Paris and a broader review of the entrepreneurial literature find women running businesses across a wide range of trades (Baijot and Le Chapelain, 2022; Craig, 2016, 2020).

The legal regime created friction, not prohibition. Chanteux (2024), studying the full population of French female patentees from 1791 to 1900, finds that 79 percent filed their patents individually and that married women were the most represented group, filing 46 percent of all female patents despite their legal subordination under the Civil Code. Marriage could have constrained women's patenting through legal and family obligations, but could also have provided capital, commercial contacts, and access to family enterprises. Beginning in the 1880s, a sequence of reforms gave women progressively greater financial autonomy, culminating in the 1907 law granting married women control over their own earnings (Lewis, 1980; McMillan, 2000).

2.2 What shaped the direction of invention?

Women in nineteenth-century France did not participate uniformly across the economy. Data from the *Enquête Industrielle* of 1860–65 show that women represented more than half of workers in the Clothing and Textiles sectors. Textiles alone employed over 40 percent of all industrial workers in France; together with food processing, these two most productive sectors accounted for 62 percent of total industrial production value in 1860 (Merouani and Perrin, 2026b). This concentration reflected the gendered division of labor, legal constraints on married women's economic activity, social norms, and the structure of local industry (Tilly and Scott, 1989). Strong occupational segregation characterized nineteenth-century France: Grantham (2012), exploiting the 1851 French census, shows that women dominated textiles and clothing trades in rural northern France while men dominated construction, agricultural crafts, and transport. Similar patterns hold for industrial Britain (Burnette, 2011).

The segregation was demand-driven. You (2020), exploiting the complete 1881 census of England and Wales, finds that the industry mix of a woman's parish mattered more than her family circumstances: married women in cotton districts worked at seven times the rate of those in agricultural areas, and even among women in the same life stage whose husbands held the same occupation, participation differed sharply between cotton and mining parishes. Boter and Woltjer (2020) reach the same conclusion for the Netherlands, showing that sectoral shifts in local labor markets account for a large share of the decline in Dutch unmarried women's participation between 1812 and 1929, and Boter (2021) confirms that local labor market structure was the dominant force shaping unmarried women's employment decisions.

Two channels could have shaped the sectors in which women patented. The first is gendered occupational knowledge. Women could have acquired sector-specific practical knowledge through their own employment, through household production, or through family enterprises. These routes differ in institutional detail but share a common logic: each gave women hands-on experience with the materials and processes of a specific industry. Arrow (1962) proposed that workers learn to improve their processes through the act of production itself; Mokyr (2005) adds that much of the knowledge embedded in techniques is tacit and transmitted through direct contact rather than written codification. Mokyr distinguishes between the epistemic base, the scientific understanding of why a technique works, and competence, the practical ability to execute it. Most practitioners needed only competence, which could be acquired through sustained engagement with production rather than through formal schooling.

I take from them a narrower claim: that such hands-on engagement with a sector could have given women sector-specific practical knowledge that positioned them to develop patentable improvements. Khan (2016) shows that women's commercial efforts were enhanced by family firms, and Merouani (2026) finds that 35.8 percent of women's collaborative ties involved a surname match indicating a family connection, versus 8.8 percent for men. A woman working in her family's textile enterprise still accumulated textile-specific knowledge; the family firm was a route through which gendered occupational exposure operated, not an alternative to it.

The depth of this engagement shows in the patent record: Khan (2016) finds that wives and widows in manufacturing firms were repeat patentees more often than other women, and Merouani and Perrin (2026a) extend the pattern to the full population of nineteenth-century French women inventors, who filed 16.8 percent more patents when affiliated with a firm.

If this channel dominates, the female employment share of local sectors, not their overall size, should predict the direction of women's patents. A department with many women in textiles should produce women who patent in textiles; a department where textiles employed mostly men should not, even if the sector was large.

The second channel is agglomeration. Bell et al. (2019) show that children who grow up exposed to more inventors are themselves more likely to become inventors; Bell et al. (2016), using the same data, demonstrate that the association runs through specific technology classes rather than innovation in general. If the same logic applied in nineteenth-century France, proximity to a large industry should have been enough to spur invention in it regardless of workforce composition. Total employment should predict women's patenting, and the female employment share should add nothing once industry size is controlled for. Living in a department with a large textile sector meant encountering textile products, techniques, and practitioners in daily life, even for women who did not work in textiles themselves.

2.3 What shaped the timing of patenting?

Most women in the patent database invented only once (Merouani, 2026; Merouani and Perrin, 2026b). The relevant question is therefore not how inventive output varied over a career but at what point in the life course a woman filed her first patent. Several forces may have shaped that timing. Filing fees, market opportunities, and the moment at which an invention became commercially valuable all could have mattered. I focus on two: human capital accumulation, and marriage and childbearing.

The first force is human capital accumulation: the years required to acquire sector-specific know-how, learn the patenting procedures, and gain access to the information networks needed to file a patent. French industrialization relied more heavily on artisanal production, small-scale manufacturing, and workshop-based innovation than on the factory system (Landes, 1969; O'Brien and Keyder, 1978). Invention in this context plausibly required what Mokyr (2005) calls prescriptive knowledge: the practical know-how for working with specific materials and processes, much of it tacit and transmitted through direct contact. If the relevant knowledge was sector-specific and largely tacit, acquiring it would have taken years of hands-on experience regardless of marital status. The time required should have varied across sectors if knowledge requirements differed: sectors where domestic skills provided a head start may have required fewer years than sectors demanding more formal or specialized knowledge.

Filing a patent also required knowledge of procedures and access to information about the centralized system (Galvez-Behar, 2019); Merouani (2026) shows that women faced an access barrier to collaborative networks, with 76.5 percent of women inventors being network isolates compared to 68.8 percent of men. Occupational knowledge and network access together set a minimum age for first patenting that applied regardless of family status.

A mid-life peak in inventive output is not peculiar to any single country or period. Sarada et al. (2019), linking American patent grantees to census records from 1870 to 1940, find that inventors were persistently about 41 years old on average, a relative age premium that remained stable for seven decades. By contrast, Jung and Ejermo (2014), using population-wide data on Swedish inventors filing at the European Patent Office between 1985 and 2007, find an overall average age in the early forties that varies by technology and period.

The second force is marriage and childbearing: the life-course demands that could delay or interrupt a woman's path to a first patent. In the 1851 French census, married women with children under three had a labor force participation rate of 60 percent, compared with 68 percent for other married women of the same age (Grantham, 2012). Twentieth-century French survey data capture the interruption pattern directly: three-quarters of married women aged 40 to 44 who had participated in the labor force had stepped away from their professional lives for more than a year (Riboud, 1985). Goldin (2006) distinguishes "jobs" defined by current circumstances from "careers" defined by long-run identity and human capital investment. She identifies children as the most important factor related to women's out-of-work spells in the former. I apply that distinction to the nineteenth-century women in my sample: child-rearing cut into the time they could devote to paid work and invention.

Humphries and Weisdorf (2015), constructing two long-run wage series for English women from 1260 to 1850, show that the gap widened specifically when industrialization moved work from flexible, home-based arrangements into rigid schedules that demanded continuous hours outside the home. Women who could commit to annual service contracts as farm or domestic servants, typically young and unmarried, kept pace with unskilled men; women confined to casual day-work compatible with childcare fell permanently behind after 1750. The wage gap was driven by the structure of work, not by a loss of skill.

Burnette and Stanfors (2012) find that in the Swedish cigar industry of 1898, mothers earned six percent higher hourly wages than childless women. Cigar-making skills were portable across employers, so continuous tenure at one firm earned no premium; a woman who left to raise children could return to any factory without losing earning power. Stanfors et al. (2014), using the same data but comparing women to men, show that career breaks did reduce women's accumulated experience, but that the cost depended on the pay system: among piece-rate workers, where individual output was directly measurable, differences in experience and job assignment explained the entire gender wage gap; among time-rate workers, where employers set wages by custom, discrimination accounted for the rest. Where work allowed skill to be demonstrated individually, career interruptions consumed time without permanently eroding capability. Humphries and Schneider (2019), drawing on records of English spinners in putting-out networks, spinning schools, and parish accounts, find the same pattern in spinning: married and widowed spinners produced within 25 percent of single spinners, a gap small enough to indicate that married women, driven by economic necessity, worked near full-time.

Kleven et al. (2019) establish the child penalty as a general feature of women's economic lives using Danish administrative data from 1980 to 2013. Women's earnings fall sharply at the first birth and stabilize at a 19 percent long-run penalty that persists for at least twenty years, grows by 7 to 10 percentage points per additional child, and holds after controlling for education. Women whose own mothers worked more relative to their fathers face substantially smaller penalties, suggesting that preferences over family and career are shaped in childhood.

Studies of career scientists corroborate the timing mechanism: Kim and Moser (2025), studying US academic scientists active in 1956 with publications tracked from 1900 to 1970, show that scientist-mothers experience a publication decline lasting roughly seven years, beginning around age 32 when the median first child would have been about one year old, with recovery in the early forties. The tenure gap is large, with only 27 percent of mothers achieving tenure compared with 46 percent of women without children. Publication declines grow with family size, reaching 54 percent for mothers of three or more children versus 37 percent for those with one or two. The timing pattern, a concentrated disruption during the early-childhood years followed by recovery, is consistent with what Grantham (2012) documents for working women in mid-century France. Whittington (2011), studying US doctorate-level scientists from 1990 to 2001, finds a motherhood penalty on the probability of patenting in academia but not in industry. The penalty disappears once prior patenting experience is controlled for, suggesting that the barrier was access to the commercial process rather than a loss of productive capacity.

What happened after childbearing ended is less settled. Tilly and Scott (1989), studying nineteenth-century textile towns in England and France, describe a substitution in which children entered the workforce and mothers withdrew to manage the household. You (2020) finds the opposite: mothers intensified their labor force participation as children reached working age, because working-class families sought to use as many laborers as possible to enlarge the household budget. Which pattern characterized nineteenth-century French women inventors has not been established.

Marriage was not purely a source of delay. It could also provide resources that offset delay, particularly through family-firm association and the commercial networks that came with it (Craig, 2020; Khan, 2016). Whether marriage on balance delayed or helped a first patent is an empirical question. The answer may also have depended on where a woman lived. Zijdeman et al. (2014), using a sample of over 53,000 marriage records from the TRA genealogical project[2] spanning 95 French departments between 1860 and 1986, find that about 60 percent of brides reported an occupation at marriage between 1860 and 1950, with persistent regional differences. Perrin (2022a), constructing a gender gap index for 86 French departments in the 1850s, confirms that departments with higher shares of women in industry displayed significantly greater gender equality. Departments where more women were commercially active offered more examples of women combining productive work with family life.

2.4 Two tests

These mechanisms make distinct predictions that the linked data can evaluate.

The direction of invention (the technology classes women chose to patent in)**.** The two channels imply different predictions about what feature of local industry matters. Gendered occupational knowledge predicts that the female employment share of local sectors, not their overall size, should determine where women patented. Agglomeration predicts the opposite: total industry size should suffice, and the gender composition of the workforce should add nothing.

The timing of patenting. The child penalty and local economic conditions make different predictions about what determines when women filed their first patent. If the child penalty operated mainly through household-level time demands, the age gap between married and unmarried inventors should be similar regardless of where they lived. If instead local conditions reduced the penalty, the gap should have been smaller in departments where more women were commercially active, because the norms, networks, and infrastructure that supported women's work may have shortened the

[2] The TRA is a French historical demography survey that traces the descendants of all men in mainland France whose surnames begin with the letters T, R, A and who married between 1803 and 1832; the initials designate the surname filter, not an acronym.

delay from child-rearing to first patent. The linked data can also distinguish between two accounts of what happened after the childbearing years. If Tilly and Scott's substitution (1989) holds, the timing of a first patent should not track the end of childbearing. If instead You's intensification (2020) holds, married women should patent predominantly after their last recorded birth.

3 Linking inventors to demographic records

The French patent record tells us who filed but not who these women were. It gives a name, an address, and a marital status at the moment of filing. It does not tell us when the inventor was born, where she came from, whom she married, or when she died. To my knowledge, no prior study has recovered this information at scale for historical female inventors in France. The country lacks a digitized national civil registry for this period, and existing work on women's invention stops at the patent document itself and some qualitative work (Chanteux, 2023; Merouani and Perrin, 2024).

I close this gap. I build, to my knowledge, the first individual-level demographic dataset for women who patented in France between 1791 and 1900. I take the 4,165 women identified by Merouani (2026) from the French Patent Database (Merouani and Perrin, 2026b),[3] and match each to collaborative genealogical databases that aggregate millions of birth, marriage, and death records transcribed by archivists and private individuals from departmental archives. The dataset covers birth year, marriage timing, geographic origin, and death year.

3.1 Search strategy and data collection

Not all inventors are equally traceable. Married women with a recorded maiden name (1,286, 30.9 percent) are far easier to locate than single women, because French naming conventions provide more information: a married woman typically appears in records with both her maiden name and her marital surname, and the fact of marriage narrows the temporal window. Widows provide even more information, since both a marriage and a husband's death must precede the patent. Single women (1,202, 28.9 percent) appear with only a birth surname and given name, offering fewer handholds for disambiguation. A third group, the largest (1,677, 40.3 percent), falls between these extremes: married women whose patent records provide a marital surname but no maiden name, because the applicant did not supply one on the patent document.

I build three separate search strategies, each tailored to the identifying information available, and tighten acceptance thresholds as the information weakens. This graduated design is the paper's main methodological contribution: a uniform rule would either admit false positives among poorly identified women or, if set strict enough to prevent them, discard many valid links. Of the 4,165 inventors, 1,212 have the three pieces needed for the most precise searching: a marital surname, a maiden name, and a first name. For these women, I construct targeted search URLs encoding the maiden name, given name, and spouse surname as standardized parameters. For single women, I search using only the birth surname and given name. For married women without a maiden name, I search using the marital surname and given name. I save the HTML of each search results page and parse it to extract names, event dates and places, source types, and relationship context. The full URL specification and parsed field list appear in Section A.1.

3.2 Linking algorithm

A single search may return dozens or hundreds of results: different people with similar names, or different records about the same person recorded inconsistently across archival sources. I solve this through a four-stage process, shown in Figure 1. I describe the process as it applies to married women with maiden names, the group with the richest identifying information.

I first eliminate implausible records using hard criteria: minimum name similarity, age-at-patenting bounds (14 to 85), and death-before-patent checks. Reported ages below 14 or above 85 are almost always linkage errors rather than true filings, because a child younger than 14 is unlikely to have filed a patent under her own name, and an inventor older than 85 would lie at the extreme upper

[3] The full database contains roughly 4,300 female patenting entities. I exclude approximately 150 family firms (e.g., "Nicod veuve et fils") that appear under a woman's name but represent a collective enterprise rather than an individual inventor.

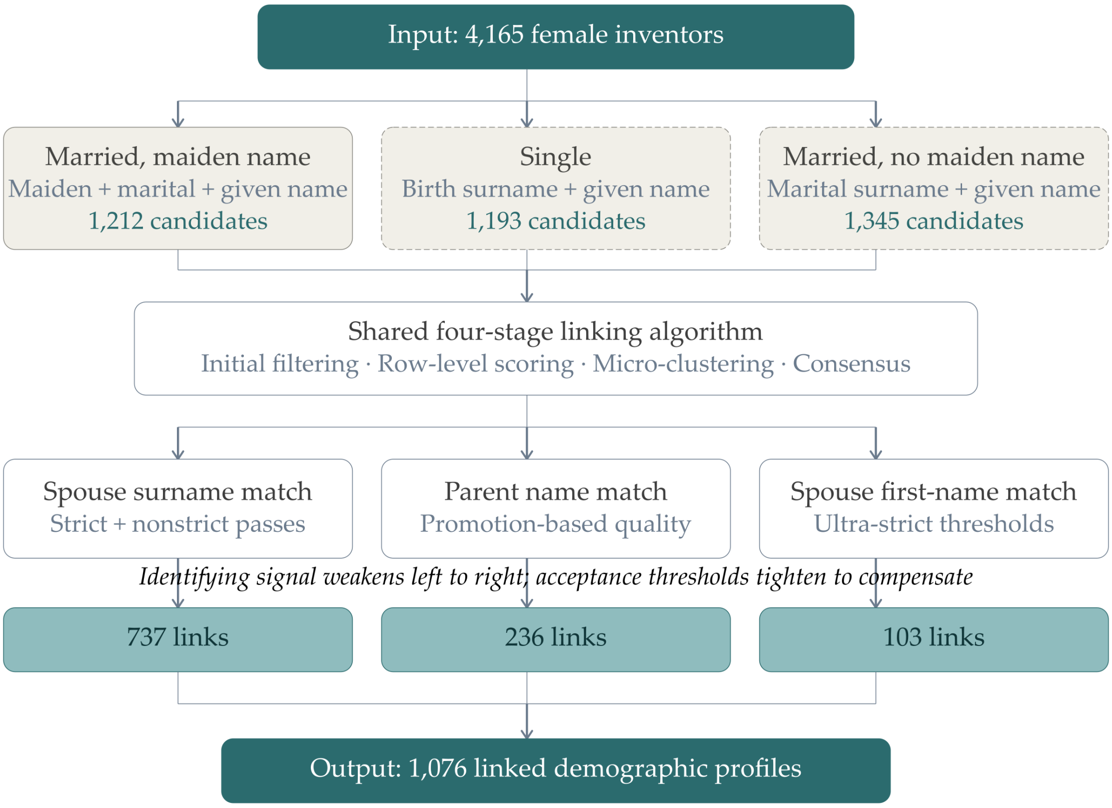


Figure 1: Linking pipeline. Of 4,165 female inventors, 415 are excluded because the patent record lists no given name. The remaining 3,750 enter one of three algorithms tailored to the identifying information available. Each algorithm filters, scores, clusters, and votes on candidate matches to produce a single best-fit demographic profile per inventor.

tail of nineteenth-century longevity. I flag records from unreliable sources but do not immediately discard them, because they may provide useful context at later stages.

Each surviving record receives a quality score combining five dimensions: name match, spouse match, place consistency, source reliability, and temporal plausibility. Each ranges from near-zero to 1.0. A multiplicative model combines them, ensuring that a record must score well on all dimensions to receive meaningful weight.

Records about the same person may appear across multiple archival sources with slight variations. I group compatible records into clusters via union-find, blocking by spouse surname and enforcing temporal agreement within tolerance windows. Two records are compatible if their birth, marriage, and death years either agree within narrow tolerances or at least one is missing, so that records with partial information can still contribute to a cluster. Each cluster represents a distinct biographical hypothesis about who the inventor was.

For each cluster, I compute a consensus biography through weighted hierarchical voting: reliable sources first for years, year-consistent records first for places. I apply plausibility vetoes to the consensus and rank surviving clusters by total weight. Results are classified as high quality, medium quality, or ambiguous based on internal agreement and the gap between competing clusters. The full algorithmic specification, including filtering thresholds, scoring components, clustering mechanics, and consensus rules, appears in Section A.2.

I run the married-with-maiden algorithm (leftmost branch of Figure 1) in two sequential passes. A strict pass using exact-name searches produces 676 links. A second nonstrict pass, restricted to inventors with recorded maiden names, captures 61 more by allowing name variants. The married-with-maiden total is 737.

3.3 Extending to single and no-maiden women

Single women present a harder linking problem because they appear with only a birth surname and given name. I adapt the same four-stage architecture but replace the spouse signal with parent names as the primary genealogical anchor (middle branch of Figure 1). Parent names serve the same disambiguation role: a record naming the same father and mother as other records in the cluster is far more likely to describe the same person, because parent-name coincidences across unrelated individuals are rare. Because the algorithm lacks an independent spouse signal, every link must earn promotion through at least one independent signal beyond the base score, such as a rare name, a geographic match, or records from official civil registers. Only promoted links are retained, all classified as high quality. This yields 236 linked single inventors.

Some married inventors lack a recorded maiden name. Searching with only the marital surname creates a circularity: every record returned has a husband with that surname because the surname was the search term. The husband's given name, however, was not part of the search, so agreement on it across records is genuine evidence. For Marie Thibaulot, I retain a cluster of 43 candidate records, all of which name her husband Augustin. I treat that full agreement as strong evidence that the records describe one woman rather than several women sharing the same marital surname. I require 70 percent agreement on the spouse's given name and apply the strictest thresholds of the three algorithms (rightmost branch of Figure 1). All 103 surviving links are high quality. Full adaptation details for both groups appear in Section A.3.

3.4 Sample composition and validation

The three algorithms together produce 1,076 linked demographic profiles: 737 married women with maiden names (676 strict, 61 nonstrict), 236 single women, and 103 married women without maiden names. Of these, 86 percent are high quality. The guiding principle is that acceptance thresholds tighten as identifying information weakens: the less I know from the patent record, the more I demand from the genealogical evidence.

I assess accuracy through stratified manual verification of 60 inventors covering both married and single women, oversampling difficult cases using a risk score that combines eight signals of matching difficulty. I independently verify each sampled inventor's dates against original document scans in departmental archives and patent filings, reading the handwritten records directly rather than relying on transcribed metadata. The full sampling procedure and estimator appear in Section A.4.

The algorithm performs well. Accuracy reaches 96 percent for birth years, 96 percent for marriage years, and 90 percent for death years, using a two-year tolerance, standard in historical demography.[4] These rates understate accuracy in the full sample, because the validation design deliberately oversamples the hardest cases. Of 50 birth comparisons, 2 exceed this tolerance, and the largest discrepancy (14 years) reflects an error in the transcribed source data rather than a near miss. Marriage dates prove easiest to verify, likely because French marriage records are well-preserved and spouses provide a strong disambiguating signal.

The 1,076 linked inventors are a selected group, but the selection is subtler than one might expect. At the patent level, Merouani and Perrin (2026b) find that 45 percent of women patentees were married, with single and widowed women in roughly equal shares of about 27 percent each, close to the population distribution of marital status in nineteenth-century France. After deduplication, the proportions are 47 percent married, 30 percent single, and 23 percent widowed. Classifying each linked inventor by the status on her first patent gives 51 percent married, 25 percent widowed, and 24 percent single.

Throughout the paper, I group inventors as *ever-married* if any patent in their record lists them as married or widowed, and as *single* if all patents record them as *demoiselle*. Twenty-six women filed patents under more than one status (e.g., as *demoiselle* on an early patent and *dame* on a later one); the first-patent percentages above count them as single, but the ever-married grouping includes them because their patent record confirms a marriage. This gives 840 ever-married and 236 single inventors. All analytical comparisons between married and single women use this ever-married definition unless otherwise noted.

Even for linked women, demographic data are incomplete. Birth years are available for 934 of 1,076 (86.8 percent), marriage years for 774 of 840 ever-married women (92.1 percent), and death years for only 576 (53.5 percent). The scarcity of death information likely reflects how genealogical databases are built. Charpentier and Gallic (2020) show that collaborative genealogies are constructed primarily through ascending research, in which genealogists trace lineages backward from a known individual; this method systematically favors direct-line ancestors and underrepresents collateral branches, women, and infant deaths. Birth and marriage records establish lineage connections; death records do not, and are less systematically recorded as a result.

The linked samples could mislead if the linking process pulled disproportionately from certain industries. I check for this by comparing the sectoral and occupational composition of each linked sample to its reference population; the full distributions appear in Figure A.1 and Figure A.3.

Table 1 reports the results for ever-married women. Of the 20 sector-level comparisons, 19 fail to reject equality at the 5 percent level. The one exception is Office Articles ($p = 0.048$). The omnibus chi-squared statistic is 26.9 with 19 degrees of freedom ($p = 0.107$), so I cannot reject the null that the two distributions are the same. Given 20 row-level comparisons, one rejection at the 5 percent level is not by itself surprising, so the omnibus test is more informative than any single row. Table 2 shows that the linked single sample tracks its reference population even more closely: no individual sector difference reaches significance at the 10 percent level, and the omnibus chi-squared is 11.3 ($p = 0.914$). Occupational composition tells a similar story, with modest shifts concentrated in the smaller categories; the full skill and HISCO distributions appear in Figure A.3 and Figure A.4. The complete distributional evidence behind these balance tests is in Section A.5.

The linking process did not detectably distort the sectoral composition of either sample. The few deviations, Office Articles in the married sample and the Elite share in the small single sample, should be kept in mind but are not large enough to undermine comparisons across the major technological or skill categories. The balance test covers sectors and occupations but not geography or birth cohort. Women with common surnames or from departments with sparser genealogical coverage are likely underrepresented in the linked sample, because name ambiguity and record scarcity make linking harder for precisely these groups.

The contribution is a new dataset and a linkage method for studying French women inventors, a population whose demographic histories have not been available at scale. I recover these histories without ready digitized census microdata for research, and the same logic can apply to any country where patent records identify names and genealogical databases preserve vital records. Differences

[4] 95 percent Clopper-Pearson exact binomial confidence intervals: birth [86, 100], marriage [84, 99], death [73, 98].

Table 1: Sectoral balance: ever-married or widowed women (N = 2,963) versus linked ever-married women (N = 840). Diff. is the linked share minus the reference share in percentage points. P-values from two-proportion z-tests. The chi-squared test at the bottom tests equality of the full 20-sector distribution.

	Ever married/widowed ($N = 2{,}963$)	Linked married ($N = 840$)	Diff.	p-value
Agriculture	2.46	2.86	+0.39	0.523
Food Processing	5.06	4.76	-0.30	0.724
Railways	0.88	0.48	-0.40	0.246
Textiles	10.26	11.55	+1.29	0.283
Machinery	7.15	8.33	+1.18	0.250
Marine & Navigation	1.01	0.36	-0.66	0.071*
Construction	3.64	3.10	-0.55	0.446
Mining & Metallurgy	1.96	1.90	-0.05	0.922
Domestic Economy	5.43	3.93	-1.51	0.080*
Road Transport	3.14	2.26	-0.88	0.185
Weaponry	0.57	0.71	+0.14	0.643
Precision Instruments	3.54	4.29	+0.74	0.315
Ceramics	1.92	2.50	+0.58	0.298
Chemistry	7.86	8.69	+0.83	0.437
Lighting & Heating	5.77	4.64	-1.13	0.206
Clothing	20.49	19.29	-1.20	0.445
Industrial Arts	4.25	4.29	+0.03	0.966
Office Articles	3.91	5.48	+1.56	0.048**
Medicine & Hygiene	5.16	4.76	-0.40	0.640
Paris Articles & Misc.	5.50	5.83	+0.33	0.711
χ^2 test (19 d.f.) = 26.92				0.107

Notes: Each cell shows the share (%) of women in the group whose first patent falls in the given sector. "Diff." is linked minus reference. p-values are from two-proportion z-tests (row-level) and a χ^2 goodness-of-fit test (omnibus). $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2: Sectoral balance: single women (N = 1,202) versus linked single women (N = 236). Diff. is the linked share minus the reference share in percentage points. P-values from two-proportion z-tests. The chi-squared test at the bottom tests equality of the full 20-sector distribution.

	Single ($N = 1{,}202$)	Linked single ($N = 236$)	Diff.	p-value
Agriculture	1.66	1.69	+0.03	0.973
Food Processing	3.74	4.24	+0.49	0.718
Railways	1.50	1.69	+0.20	0.821
Textiles	9.90	13.14	+3.24	0.137
Machinery	5.57	5.51	-0.07	0.968
Marine & Navigation	0.67	0.85	+0.18	0.759
Construction	3.49	4.66	+1.17	0.384
Mining & Metallurgy	1.16	1.27	+0.11	0.890
Domestic Economy	5.99	4.24	-1.75	0.288
Road Transport	3.66	2.54	-1.12	0.391
Weaponry	0.67	0.42	-0.24	0.667
Precision Instruments	2.91	2.12	-0.79	0.498
Ceramics	1.66	2.12	+0.45	0.625
Chemistry	5.99	8.05	+2.06	0.235
Lighting & Heating	5.07	4.66	-0.41	0.790
Clothing	24.96	25.42	+0.47	0.880
Industrial Arts	5.99	5.93	-0.06	0.973
Office Articles	4.41	3.81	-0.60	0.680
Medicine & Hygiene	4.74	2.97	-1.78	0.226
Paris Articles & Misc.	6.24	4.66	-1.58	0.350
χ^2 test (19 d.f.) = 11.28				0.914

Notes: Each cell shows the share (%) of women in the group whose first patent falls in the given sector. "Diff." is linked minus reference. p-values are from two-proportion z-tests (row-level) and a χ^2 goodness-of-fit test (omnibus). $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

in naming conventions and record quality will affect feasibility, because the algorithm calibrates what it demands to what each inventor's record provides.

4 Geographic origins

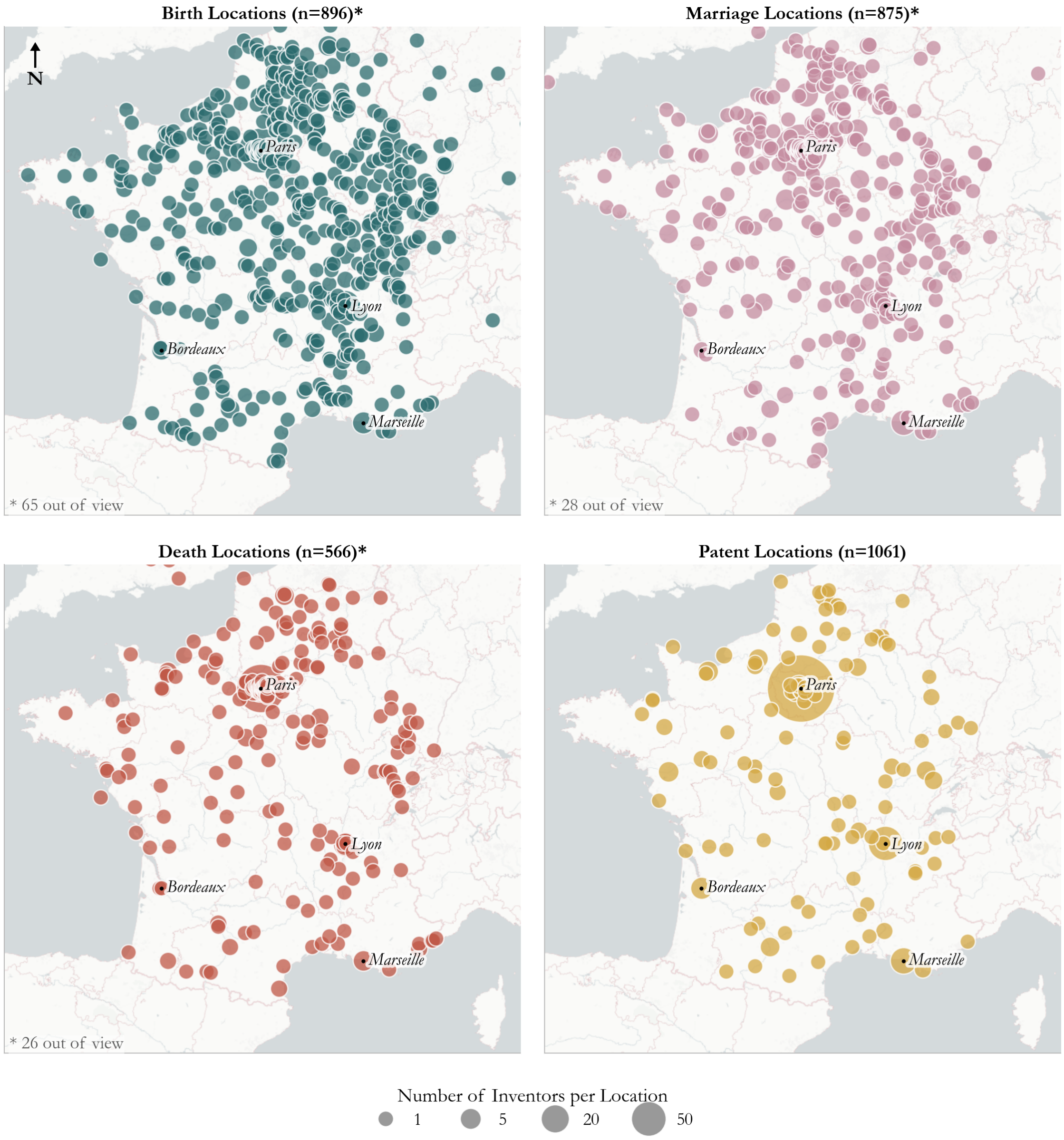


Figure 2: Spatial patterns of birth, marriage, death, and patent locations among linked French women inventors, 1791-1900

4.1 Patent addresses and residential locations

French patent records required an address but no proof of domicile (Loi du 5 juillet 1844, art. 5-6).[5] If these addresses reflect where inventors lived, they are valuable evidence about the geography of innovation. If they reflect administrative convenience, they distort the picture. In my linked sample, 80 percent of patent addresses fall in Paris. Studies of Italy and the United States find geographic concentration in patenting, but it is dispersed across regions or states rather than funneled through a single city (Akcigit et al., 2017; Nuvolari and Vasta, 2017). The French concentration in Paris

[5]Patent addresses are often taken at face value in studies of the geography of innovation, even with known biases due to assignee-provided addresses.

Table 3: Distance from marriage location to patent address by group

Group	N	Mean (km)	Median (km)	% $\leq$ 50km	% > 200km
Parisian Residents	333	1.0	0.0	100.0%	0.0%
Paris-Address Non-Parisians	247	247.9	212.7	6.1%	56.3%
Non-Paris Residents	136	66.6	1.6	69.9%	11.0%

Notes: Distances calculated using geodesic formula. Sample restricted to married inventors with valid marriage and patent coordinates.

is an artifact of centralized filing, not evidence of where inventors lived. The civil records tell a different story: births are widely distributed across France, and only 47 percent of marriages occurred in Paris. I test this divergence directly by comparing marriage locations with patent addresses for 727 married inventors.

I stratify inventors based on where they married and where they filed patents. This produces four groups:

- Parisian Residents (333 inventors, 46 percent): Married in Paris, patent address in Paris
- Outward Migrants (11 inventors, 2 percent): Married in Paris, patent address outside Paris
- Paris-Address Non-Parisians (247 inventors, 34 percent): Married outside Paris, patent address in Paris
- Non-Paris Residents (136 inventors, 19 percent): Married outside Paris, patent address outside Paris

Table 3 shows the distances between marriage and patent locations for each group. Parisian Residents have a median distance of 0.0 km. Non-Paris Residents have a median of 1.6 km, with 69.9 percent within 50 km of their marriage location. Both patterns are consistent with residential addresses. Paris-Address Non-Parisians show a median of 212.7 km and a mean of 247.9 km. This gap is the key evidence: women who married outside Paris but listed Paris patent addresses had married, on average, hundreds of kilometres from where they filed (see Section B.1 for the full forensic analysis, including Figure B.1, Table B.1, and Table B.2).

Patent agents explain part of this pattern. Among inventors married outside Paris, 77.6 percent of those with listed agents have Paris addresses, compared to 50.8 percent of those without agents (difference: 26.7 percentage points, $p < 0.001$). Under nineteenth-century French patent law, applications were processed through a central office in Paris and could be submitted by a *mandataire* whose address appeared in the records (Galvez-Behar, 2019). This institutional centralization contrasts with Italy, where applications were lodged via local prefetture and patents were assigned using inventors' stated residences (Nuvolari and Vasta, 2017). Death records suggest that at least 8.1 percent of inventors may have used Paris addresses without permanently residing there (detailed evidence in Section B.1).

Addresses outside Paris appear more reliable, and Figure 2 shows why this matters. Birth and marriage locations reveal geographic dispersion that patent addresses hide. Lyon, Bordeaux, and Marseille remain on the patent map but shrink relative to Paris. The many smaller cities that dot the civil record maps vanish entirely.[6] For the analyses that follow, I use birth departments rather than patent addresses to locate inventors' geographic origins.

4.2 The North-South gradient

The birth maps in Figure 2 hint at a feature of nineteenth-century French economic geography: the divide between the more developed North-East and the agricultural South-West, first mapped by Dupin (1826) and later refined by Braudel (1984, pp. 337–43). Nuvolari et al. (2023) use the Braudel line (Rouen to Geneva) to characterize the spatial concentration of French patenting between 1791 and 1844. A dummy for departments north of this line is a strong predictor of women's inventor rates in my data ($p < 0.001$, $R^2 = 0.28$), and the Dupin line (St. Malo to Geneva) performs similarly ($R^2 = 0.29$; see Table B.3 in the appendix). Women inventors came disproportionately from the more developed half of France, a pattern consistent with the broader economic geography

[6] Figure B.2 traces the full geographic flows across birth, marriage, and death.

of the period. But this aggregate geographic gradient is a blunt instrument. Whether local industry shaped the *direction* of women's inventive activity requires a different approach.

5 Local industry and the direction of invention

The previous sections establish that women inventors' origins were geographically dispersed, that patent addresses overstate Paris concentration, and that rates were still higher in the more developed North-East than in the agricultural South-West (Table B.3). These findings describe where women came from. They do not explain whether the local economy also predicted the *direction* of inventive activity. I now turn to this question.

Women's patents span twenty technology classes, including textiles, clothing, chemistry, machinery, and domestic goods. Did a woman from a textile department patent in textiles? Did a woman from a chemistry department patent in chemistry? If so, the local economy did not merely produce inventors; it also predicted what they invented. The hypothesized channel is local industrial knowledge. A woman who spent years working with specific materials and processes could develop the practical understanding needed to improve them, even without formal schooling. An alternative is pure agglomeration: departments with large industries produce more inventors in those industries regardless of gender composition. The empirical challenge is to separate the two.

I test this using the 1860 department-level industrial survey, which records female and male employment shares across sixteen manufacturing sectors for all ninety metropolitan departments. I construct a concordance mapping each of the twenty patent technology classes to the most closely corresponding of the sixteen 1860 industries, then ask whether inventors systematically patent in industries that are concentrated in their birth departments.

5.1 The technology-class match

The simplest test asks whether inventors tend to come from departments that specialize in their patent technology class, without imposing any model on how they choose among industries. If no such concentration exists, there is nothing to decompose.

The concordance has three levels of certainty: eight technology classes where the mapping is unambiguous (Clothing maps to the survey Clothing industry, Textiles to Textiles, Chemistry to Chemicals, and so on), three technology classes that map directly but use a different variable format in the survey (Agriculture, Railways, Road Transport), and seven technology classes (273 inventors, 35 percent of the sample) where the mapping requires judgment. For these seven, the twenty major patent classes are too coarse to assign a single survey industry, so I use the finer-grained minor-class classification (98 subcategories nested hierarchically within the twenty major classes) to assign each inventor to the most appropriate 1860 industry at the individual level. For example, the Machinery major class contains both sewing-machine and metal-working-tool patents; the minor class identifies which. Seven inventors hold sewing-machine patents (minor class 05.7) and are assigned to the survey Textiles industry rather than Metallurgy; within Domestic Economy, nineteen hold furniture patents (minor class 09.4) and are assigned to Furnishings. I exclude Weaponry (six inventors) as too small for any statistical test. The full concordance, with inventor counts and rationale for each mapping, appears in Table~C.1 in the appendix.[7]

For each department d and industry s, I compute the female Location Quotient (LQ):

$$LQ^F_{d,s} = \frac{e^F_{d,s} / \sum_s e^F_{d,s}}{\sum_d e^F_{d,s} / \sum_{d,s} e^F_{d,s}}$$

where $e^F_{d,s}$ is the share of the department's female working-age population (15 to 65) employed in industry s, as recorded in the 1860 industrial survey. A value above one indicates that the

[7] This analysis requires a known birth department, which restricts the sample to the 790 inventors (73 percent of the linked sample) whose birth location could be assigned to a metropolitan French department. The remaining inventors either have no recorded birth location or were born outside metropolitan France. The restricted subsample closely resembles the full sample on marital status (80 versus 78 percent married), technology class distribution (clothing 20 versus 21 percent, textiles 12 versus 12 percent), and birth decade.

department employs women in that sector at an above-average rate. For each of the 453 inventors in the eight directly mapped technology classes with birth-department data, I look up the LQ of her birth department in her patent's concordance-mapped industry (452 have non-missing LQ values; 1 inventor is in a department with zero recorded employment in her matched industry).

A concrete case illustrates the measure. Marie Massart née Roellinger was born in Haut-Rhin, the Alsatian textile heartland where cottage spinning and weaving pervaded rural households and women moved seasonally between farm labor and textile work (Harvey, 2000), producing employment figures that exceed 100 percent of the female working-age population in the industrial survey.[8] Textiles accounts for 98 percent of Haut-Rhin's female industrial employment, compared with 71 percent nationally; the resulting LQ is 1.39. In 1889, at age 35, Massart patented a "new product called Enamel Fabric, and its manufacturing processes" (1BB199376), a textile innovation rooted in the industry that surrounded her from childhood. Her match score of 1.39 says that her birth department was 39 percent more specialized in her patent technology class than the national average.

Across all 452 inventors with valid match scores, the mean is 1.71: on average, the industry matched to an inventor's patent had a female employment share in her birth department 71 percent above that industry's national average. The median, however, is 0.61: more than half of inventors come from departments where their patent technology class is not especially concentrated, with a tail of strongly matched department-inventor pairs pulling the mean upward. A median below one is not evidence against the hypothesis. It reflects the dominance of clothing and textiles, the two largest patent classes, whose female employment was so broadly distributed across France that few departments were highly specialized in them. In contrast, industries such as chemicals or lighting, where female employment was more geographically uneven, are the ones that generate the high match scores driving the mean.

The average match score could be high by coincidence if inventors simply cluster in large departments that employ many workers in every sector. To test this, I keep every inventor in her birth department but randomly reassign which technology class she "patented" in. Under this shuffling, Massart would still come from Haut-Rhin, but might be assigned to ceramics instead of textiles; Haut-Rhin has zero women in ceramics, so her match score would drop to zero. I repeat this 10,000 times and recompute the average match score in each permutation. The observed mean (1.71) exceeds 99.2 percent of permuted means, yielding a one-sided empirical p-value of 0.0078.[9] The same test using male employment shares also rejects the null (empirical $p = 0.0094$). The full permutation results, including stratified and decomposition variants, appear in Table C.2 and Figure C.3 in the appendix. The concentration is detectable only when the concordance is clean: extending the test to all technology classes, including the seven that require judgment-call mappings, produces a non-significant result ($p = 0.575$), consistent with measurement error in the mapping rather than the absence of an underlying association.

The permutation test establishes that inventors are geographically concentrated in departments that specialize in their patent technology classes, but it cannot distinguish whether this concentration reflects female employment specifically or local industry in general. Female and male employment are 97 percent correlated across departments: a department with many women in textiles almost invariably has many men in textiles as well. Any comparison across departments faces this same collinearity.

5.2 Women invented where women worked

The conditional logit sidesteps this collinearity by changing the unit of comparison. Instead of asking whether textile departments produce more textile inventors, it asks whether, within a given department, an inventor patents in the sector where women are concentrated rather than in

[8]The 1860 *Enquête Industrielle*, conducted by the *Statistique Générale de la France* between 1860 and 1865, enumerated workers (*ouvriers*) by establishment and industry branch, recording counts and daily wages separately for men, women, and children (Chanut et al., 2000). The survey counted positions at establishments, not unique individuals. Because many women in textile regions moved seasonally between agricultural and industrial work (Harvey, 2000), the same woman could appear in multiple sectors, producing employment figures that exceed 100 percent of the female working-age population.

[9]The result does not reflect compositional differences in sector choice between single and ever-married women. Shuffling sectors within marital-status strata, rather than pooling, yields $z = 2.631$ and $p = 0.0096$, nearly identical to the unstratified test.

the sectors where they are not. This within-department comparison absorbs all department-level confounders (wealth, culture, infrastructure, overall industrialization) because they are identical for every sector choice facing the same inventor. Sub-department heterogeneity remains outside the comparison: neighborhoods of the same city could differ in schools or workshops, and rural communes could differ from urban centers within a department. The remaining variation runs across sectors within each department, where industry size and gender composition can move independently. Unlike across departments, where female and male employment rise and fall in lockstep, within a given department one sector can be large and female-dominated while another is large and male-dominated. This is the variation that separates the gendered-knowledge channel from pure agglomeration.

5.2.1 The female employment share dominates

The model treats each inventor's patent technology class as a discrete choice among the industries available in her birth department.[10] It assumes every inventor could have patented in any industry in the choice set; in practice, a woman's skills constrained her to a narrower set, and the coefficients represent the average pattern across all inventors, not a prediction about any individual inventor's choice.

For each of the 453 inventors in the eight directly mapped technology classes, I construct a choice set of twelve industry alternatives (the eight mapped industries, with Mining and Metallurgy entered as separate alternatives, plus three industries that no inventor chose but that represent plausible alternatives: Leather, Wood, and Construction) and observe the department's female and male employment rates in each sector.[11] The model estimates how these employment rates affect the probability that an inventor patents in a given industry. Because the comparison runs across sectors within the same inventor, everything about that inventor, her age, her marital status, her birth decade, the wealth of her department, drops out. Only the characteristics that differ across sectors can explain her choice.

I estimate three specifications that build from the raw female association to the preferred decomposition of size and female share (Table 4). Model (1) includes only female employment. Model (2) adds male employment to control for overall sector size. Model (3), the preferred specification, decomposes employment into total industry size and the female employment share:

$$\Pr(y_i = s \mid d_i) = \frac{\exp(\beta_1 \,\text{TotalEmp}_{d_i,s} + \beta_2 \,\text{FemShare}_{d_i,s})}{\sum_{k\in\mathcal{S}} \exp(\beta_1 \,\text{TotalEmp}_{d_i,k} + \beta_2 \,\text{FemShare}_{d_i,k})}$$

where $y_i = s$ indicates that inventor i from department d_i patents in industry s, $\text{TotalEmp}_{d,s}$ is total employment (female plus male) in industry s in department d, and $\text{FemShare}_{d,s}$ is the female share of that total. Here, $\mathcal{S}$ is the twelve-industry choice set described above, and k indexes the alternatives in that set. The denominator is the conditional-logit normalizing term: it adds the exponentiated index over all twelve possible industries so that the predicted probabilities sum to one for inventor i. The coefficient β_1 captures agglomeration: do larger industries attract more patents regardless of gender composition? The coefficient β_2 captures the gendered channel: conditional on industry size, do sectors with a higher share of female workers attract disproportionately more women's patents?

Both channels operate. Total employment is positive and significant ($\beta_1 = 0.0102$, $p < 0.001$): larger industries attract more patents. The female employment share is also positive and significant ($\beta_2 = 1.82$, $p < 0.001$): conditional on industry size, sectors with a higher share of female workers attract disproportionately more women's patents. To compare magnitudes, I standardize both

[10] The conditional logit is the natural model when the outcome is a choice among discrete alternatives rather than a continuous variable.

[11] Survey-based employment shares likely undercount women's actual economic engagement, particularly in home-based textile work where women were invisible to establishment surveys. Chiswick and Robinson (2021), reclassifying women in the 1860 United States census who lived with self-employed relatives as unreported family workers, estimate that true female labor force participation was 57 percent, more than three times the official 16 percent. If similar undercounting affected the French *Enquête Industrielle* and disproportionately affected sectors with high female employment, the measured association would be attenuated; the positive coefficient I find would then understate the true relationship.

Table 4: Conditional logit: industry choice as a function of local employment. Model (1) includes female employment only. Model (2) includes both female and male. Model (3) decomposes into total employment and the female employment share.

	(1)	(2)	(3)
Female employment	0.0246***	0.0722***	
	(0.0067)	(0.0291)	
Male employment		-0.0526***	
		(0.0293)	
Total employment			0.0102***
			(0.0024)
Feminization share			1.8217***
			(0.3429)
Log-likelihood	-1100.95	-1074.24	-1068.53
N	5436	5436	5436
Wald ($\beta_f = \beta_m$)		χ^2=4.588, p=0.0322	

Note: Conditional logit with inventor-level fixed effects. Choice set includes nine Tier-1 industries (eight patent classes, with Mining and Metallurgy as separate alternatives) plus three additional sectors (12 total). Cluster-bootstrapped standard errors in parentheses (999 replications, resampling the 86 birth departments with replacement). Significance from bootstrap percentile confidence intervals. *** p<0.01, ** p<0.05, * p<0.10. Robustness across concordance tiers, subsamples, and alternative estimators reported in Table 5.

coefficients so that a one-standard-deviation increase in each variable is comparable. On this scale, the point estimate of the female employment share (0.49) is approximately twice as large as that of total employment (0.25), though this ratio inherits uncertainty from both coefficients. The gender composition of the workforce is the dominant predictor of where women patented.[12]

Model (2) is consistent with this decomposition. The female employment coefficient is positive ($\beta = 0.0722$) while the male employment coefficient is negative ($\beta = -0.0526$). The negative male coefficient does not mean that male employment actively discourages women's invention. It reflects the fact that sectors with high male employment relative to female employment, such as mining and construction, had very few women workers in the 1860 industrial survey and produced almost no women's patents. In a sector like textiles, where both female and male employment are high, the positive female coefficient and the negative male coefficient pull in opposite directions; the net effect is positive because the female coefficient is larger in absolute value. The two coefficients are opposite in sign and statistically distinguishable from each other (bootstrapped Wald $\chi^2 = 4.6$, $p = 0.0322$): female employment attracts women's patents into a sector while male employment does not.

A concrete case illustrates the within-department comparison. Marie Josephe Thevenet née Gruyer was born in Isère, where three industries employed women: Clothing (88 percent female), Chemicals (35 percent), and Food processing (10 percent). Thevenet patented an "automatic mounting system for umbrellas, parasols, and sunshades" (1BB121740), an invention squarely in the Clothing industry where women in her department were concentrated.

5.2.2 The result survives measurement and model checks

The model asks whether this pattern holds systematically: across all inventors, did women patent in the sector that was more female within their own department? The answer is yes. The gendered channel is robust across alternative samples and concordances, though the coefficient shrinks when the concordance includes less certain technology-class mappings. I re-estimate the preferred

[12] All inventors born in the same department see the same employment figures, so the 453 inventors are not truly independent observations. The effective sample size is closer to 86 departments. To account for this, I compute standard errors in Table 4 by repeatedly resampling entire departments (999 bootstrap replications). I assess significance from bootstrap percentile confidence intervals rather than from dividing the coefficient by the bootstrap standard error, because the bootstrap distribution is skewed and the standard error, which assumes symmetry, overstates the uncertainty around coefficients whose bootstrap distributions are concentrated away from zero. This correction is severe for Models (1) and (2), where raw worker counts scale with department size: standard errors increase roughly fivefold (e.g., 0.0291 vs. 0.0067 for female employment). The correction is milder for Model (3), where standard errors roughly double. The reason is that the female employment share captures the *composition* of each sector's workforce, not its size, and this composition differs across sectors even within the same department. All coefficients remain significant after this correction; see Section C.3.

Table 5: Re-estimation of the conditional logit across sample variants. Each row is a separate estimation on the indicated sample or concordance.

	Total employment	Fem. employment share	N
8 direct sectors (baseline)	+0.0102***	+1.82***	453
+ agriculture, transport (11 sectors)	+0.0054**	+1.68***	509
+ judgment-call sectors (17 sectors)	+0.0033	+1.01***	785
Excluding Paris-born inventors	+0.0082***	+1.51***	343
Born 1840–1880 only	+0.0087**	+1.64**	191
Proximate department	+0.0124***	+1.97***	551
Mixed logit (random coeff., IIA)	+0.0128***	+2.33***	440

Note: Each row re-estimates the preferred conditional logit decomposition (Table 4, col. 3) on the indicated sample or concordance. Total employment is total sectoral employment (female plus male); Fem. employment share is the female share of that total. All rows use department-level cluster bootstrap (999 replications); significance from bootstrap percentile confidence intervals. *p<0.05; **p<0.01; ***p<0.001. Last row: mixed logit with normally distributed random coefficients (model-based inference).

decomposition from Table 4 on seven variants; Table 5 reports the coefficients from each, with department-level cluster bootstrap inference throughout.

The concordance tiers in Table 5 trace the effect of measurement error. The eight directly mapped sectors produce a female employment share coefficient of 1.82. Adding three sectors with clean but indirect survey mappings (agriculture, transport) reduces it modestly to 1.676. Adding the seven sectors that require judgment-call mappings produces a larger drop to 1.012. This decline is what measurement error produces. If the gendered channel were genuinely weaker in the added sectors, the coefficient should change regardless of mapping quality; instead, it falls in proportion to how uncertain the mapping is. I treat the eleven-sector concordance as preferred because it adds unambiguous mappings while avoiding the noise from judgment-call technology classes.

The remaining robustness checks in Table 5 confirm stability. Excluding the 199 Paris-born inventors preserves the result (N = 343). Assigning each inventor to the department of her most recent civil event before patenting, rather than her birth department, increases the female employment share coefficient from 1.82 to 1.97 (N = 551); the birth-department estimate is conservative.

The conditional logit assumes that adding or removing a sector from the choice set does not change the relative odds of choosing among the remaining sectors. This assumption is unlikely to hold when textiles and clothing, the two most feminized sectors, are closely related: a woman choosing between them is not choosing between independent alternatives in the way that a woman choosing between textiles and mining would be. Formal tests reject this independence assumption for 4 of 7 testable sectors at the 5 percent level.

I address this with a mixed logit that allows the responsiveness to female employment to vary across inventors rather than assuming all women respond identically. The mean coefficient is 2.33 ($p < 0.001$), larger than the baseline estimate of 1.82, and the spread across inventors is substantial (SD = 3.13). The mixed logit, which relaxes the independence assumption, produces a larger coefficient than the baseline. The sign and significance survive; the larger magnitude may also partly reflect the mixed logit's assumption that inventor responses follow a normal distribution. Full details appear in Section C.4.

A separate concern is temporal distance: local industrial structure is observed in 1860, but inventors were born between 1764 and 1881. A woman born in 1790 acquired her practical knowledge decades before the industrial survey, and her department's industrial structure may have shifted in the interim. I test whether the female employment share effect weakens for women whose formative years fell far from 1860. The effect does weaken mildly for distant cohorts, but the net effect at the sample's mean distance of 18 years remains close to the baseline (1.81 versus 1.82). Splitting the sample confirms this: inventors within twenty years of 1860 (N = 284) produce a coefficient of 2.01; those twenty to forty years away (N = 139) produce 1.89. Both are highly significant. 61 percent of the sample falls in the first group. The results reflect long-lasting regional patterns in which industries employed women, not the conditions of any single year (Zijdeman et al., 2014). Full details appear in Section C.5.

5.2.3 The match reflects local knowledge rather than agent sorting or migration

The sector-level breakdown (Figure C.1 in the appendix) shows which industries drive the result: Lighting, Chemicals, and Clothing show strong geographical concentration, while textiles falls below one because textile employment was so widespread nationally that no department was truly specialized in it.

The pattern extends beyond Massart's Alsatian textiles and Thevenet's Isérois accessories. In Bouches-du-Rhône, home to the Marseille soap industry where 2.2 percent of women worked in chemicals (compared with 0.2 percent nationally), Marie Louise Trouilhas née Dreveton patented a process for "manufacturing salt, soda, and potash crystals extracted from olive marc" (1BA8574) at age 47 in 1839. Olive marc was a byproduct of the region's olive oil production, and Trouilhas turned it into a patentable chemical process. Across three departments, three industries, and three women, the same logic holds: each invented in the sector where women in her locality worked.

One alternative mechanism is agent sorting: patent agents in Paris may have specialized by technology class, and if a textile agent recruited clients from textile departments, the observed match could reflect agent networks rather than inventor knowledge. The conditional logit conditions on the inventor's department, so agent sorting would need to operate at the department-sector level to produce the pattern. Excluding inventors born in Paris, where the patent office and agents were concentrated, preserves the result. A direct test also finds no relationship: department-level female industrial employment does not predict whether an inventor used a patent agent ($p = 0.35$). Women who sought agents may have differed from those who did not, but the absence of any department-level association makes it unlikely that agent networks drive the sector match.

A further test exploits inventors who moved between departments. Among the 80 who moved to Paris, sector matching is stronger in their destination than their origin (70.0 percent positive, Wilcoxon $p = 0.0001$), consistent with Paris offering high female employment across virtually every sector. The 45 non-Paris movers show no such pattern, though the subsample is too small for reliable inference. A conditional logit that includes employment shares from both the birth and marriage departments as competing predictors suggests that marriage-department employment dominates for movers, though the small number of movers makes this evidence suggestive rather than definitive (Section C.7). For women who moved, the proximate department is the more relevant unit of exposure; I use birth department for the main analysis because it is available for a larger sample, because birth and marriage department coincide for non-movers, and because long-distance moves between non-Parisian departments were uncommon in nineteenth-century France.

Occupation data, where available, confirms that at least part of the knowledge channel ran through own-work experience. Among the 70 clothing inventors for whom a profession is recorded, at least half list a directly clothing-related trade: corset makers, seamstresses, milliners. Others list occupations outside clothing or generic descriptions such as "without profession" that do not indicate industry experience. The conditional logit cannot isolate how women acquired that sector-specific knowledge, because all plausible routes would generate the same statistical pattern. But the occupation evidence establishes that women with hands-on experience in a trade patented in that trade, consistent with the knowledge mechanism regardless of how that experience was acquired. The conditional logit also cannot separate knowledge from access to production infrastructure. A woman born in a textile department had both textile expertise and proximity to textile suppliers, equipment, and distribution networks; the design identifies that gendered local industry predicted sector choice but cannot determine whether the channel ran through what women knew or what they could physically reach.

The local economy predicted which sectors women patented in. Whether it also predicted *when* they first patented is the question I turn to next.

6 Life cycle of invention and family formation

6.1 The marriage divide and life-cycle timing

Married women patented at a median age of 41.5. Single women patented at 30.0 (Figure 3). That 12-year gap raises an immediate question: did marriage and childbearing set the clock on when women invented? I trace the life-cycle patterns using 934 linked inventors for whom I observe both a birth year and the year of the first patent: 702 ever-married and 232 single; exercises that require marriage dates use only the ever-married subsample.

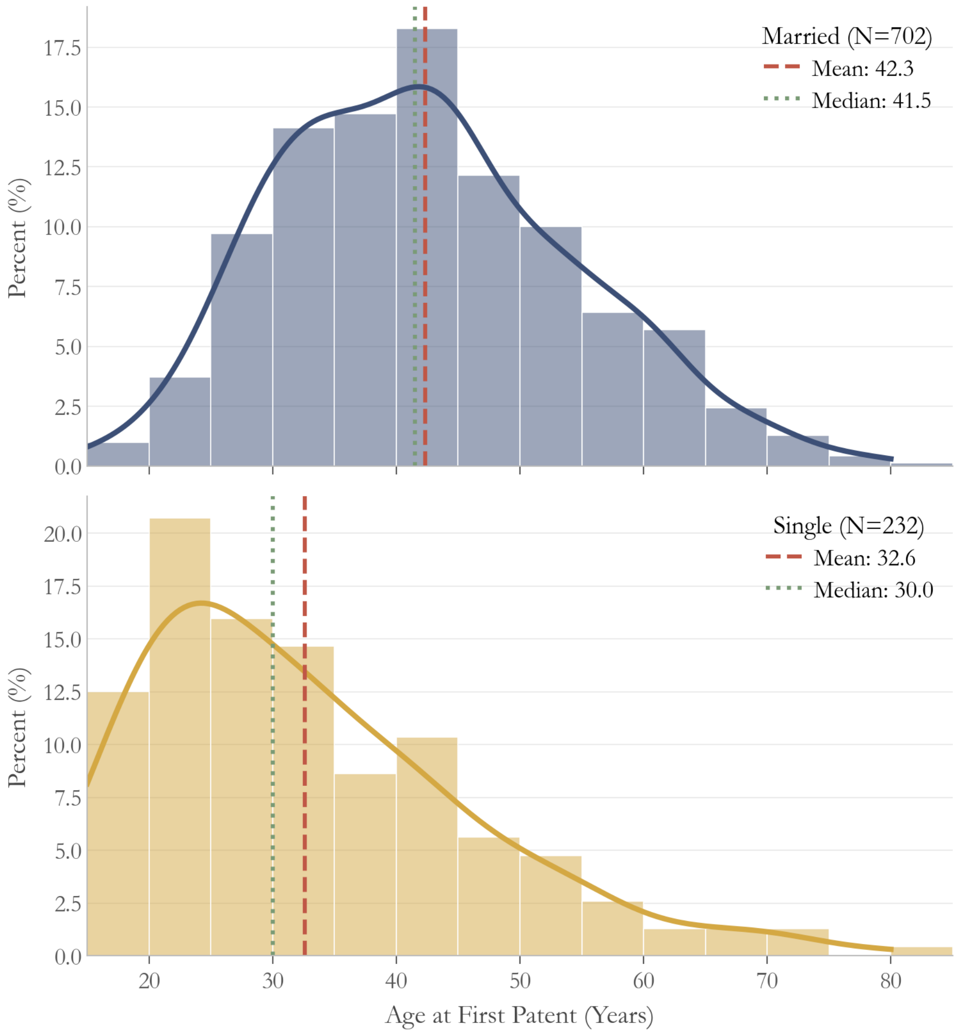


Figure 3: Age at first patent by marital status. Top panel: married or widowed inventors (N = 702, mean = 42.3, median = 41.5). Bottom panel: single inventors (N = 232, mean = 32.6, median = 30.0). Single women filed their first patent roughly 12 years younger than married women. Both panels use the same x-axis scale for direct comparison.

The two distributions barely overlap (Figure 3). Among married inventors, the interquartile range spans 33 to 50; single women cluster a full decade earlier. The gap is not a matter of outliers: the central masses occupy different decades of life entirely. Single women, unconstrained by marriage and childbearing, patented at a median age of 30.0, younger than the 35-to-55 productive window that Akcigit et al. (2017) identify for American inventors. Most married women filed in their forties. The tension between family and invention was not peculiar to women: Akcigit et al. (2017) show that American inventors also delayed marriage and had fewer children than the general population, though in their case the delay fell on family formation rather than on productive work.

6.1.1 Marriage, invention, and death

Figure 4 places first patenting in the broader sequence of major life events for married women. Among the 654 women with observed marriages, the mean age at first marriage is 24.0 years and the median is 23.0. For the 702 married women with birth and patent dates, the mean age at first patent is 42.3 and the median 41.5. Among the 438 married women for whom I observe a death year, average lifespan is 68.0 years with a median of 70.0. The medians sketch a three-stage sequence for married inventors: marriage in the early twenties, a first patent in the early forties, and

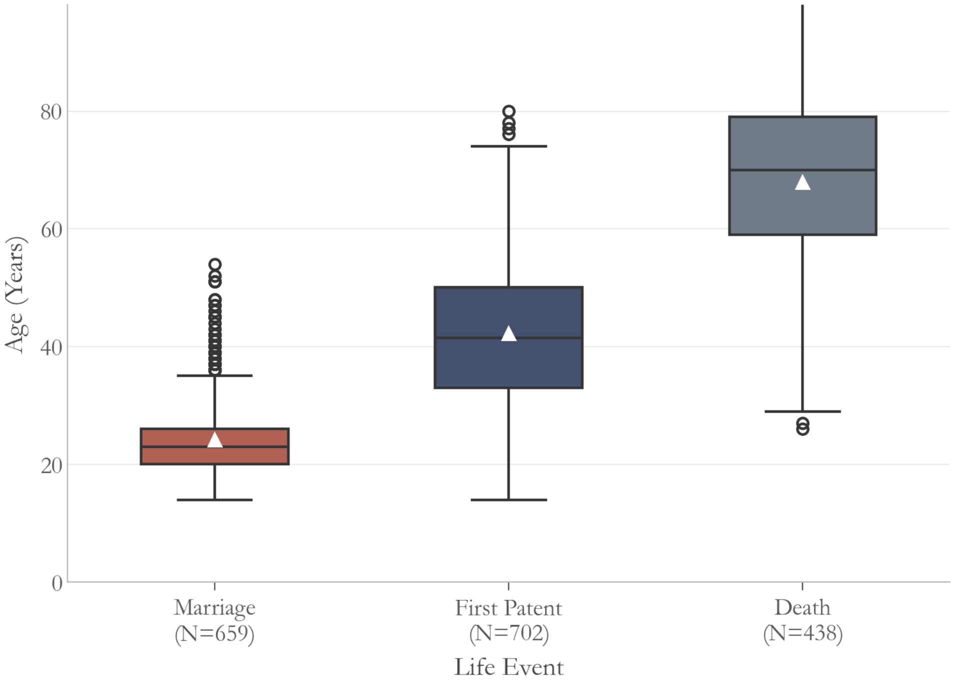


Figure 4: Biographical milestones for married inventors. Comparison of age distributions for marriage (N = 654), first patent (N = 702), and death (N = 438). Boxplots show the median (line), mean (triangle), and interquartile range. The gap between marriage (mean 24.0) and first patent (mean 42.3) averages 17.8 years.

death around age seventy. On average, women file their first patent about 17.8 years after marriage and roughly 25.8 years before death; only 4.2 percent of married women patent before marrying (Figure D.1 in the appendix shows the full distribution of marriage-to-patent lags). Invention was a mid-life activity.

The dip-and-recovery pattern that Kim and Moser (2025) document for scientist-mothers (Section 2) offers the closest parallel. The early-forties recovery in their data corresponds to the age at which married women in this sample file their first patents. Jung and Ejermo (2014) report a consistent pattern among contemporary Swedish inventors, where women patent at an average age of about 40.

The pattern matches the withdrawal-and-return sequence that Riboud (1985) documents for French married women (Section 2). The practical knowledge involved was sector-specific but not firm-specific, so a long break from work need not have eroded it, as Burnette and Stanfors (2012) show for women in the late nineteenth-century Swedish tobacco industry. The married inventors in this sample appear to follow the same sequence: work before marriage, withdraw around childbearing, and file a first patent in the early forties, roughly two decades after marriage.

In Goldin's (2006) terms, these women's intermittent attachment to production resembles "jobs" rather than "careers." But the sector-choice evidence in Section 5 points to durable sector-specific knowledge. Women patent in the industries that employed women in their birth departments. Goldin's dichotomy does not anticipate this combination.[13][14]

The mid-life average masks a split by productivity. One-time patentees (765 women) file their first patent at a mean age of 40.8; serial patentees (two or more patents, N = 169) file their first patent at 35.7, 5.1 years earlier ($p < 0.001$, Cohen's $d = 0.4$; Figure D.2, Figure D.3, Figure D.4, and Figure D.5 in the appendix). Single women were serial inventors at more than twice the rate of married women (30.2 versus 14.1 percent, $p < 0.001$). Academic mothers achieve tenure at much

[13] See Merouani and Perrin (2022) for a review of the literature on gender, work, and development during industrialization.

[14] Khan (2016) documents cases of women who acquired deep technical expertise through family firms, including metallurgy, oenology, and industrial chemistry.

lower rates than women without children (Kim and Moser, 2025). Children were associated with not only the timing of a first patent but also whether women patented more than once. Among the 101 single inventors who eventually married, 91.1 percent patented before marrying, with a mean gap of 7.6 years between the first patent and the wedding. The sequencing suggests that these women built inventive careers first and formed families second, though the correlation between single status and serial patenting could also reflect selection: women with greater inventive ambition may have delayed or forgone marriage.

6.2 The marriage gap is stable across economic environments

The gender composition of local industry predicted *what* women invented. Did it also predict *when*? If departments with more female industrial employment offered denser networks of commercially active women, or more opportunities to combine domestic duties with productive work, the marriage gap, the age difference between married and single women at first patent, might shrink in more industrialized places. The child-penalty literature predicts little response to such structural advantages (Kleven et al., 2019); narrowing the gap would require a shift in gender norms or workplace flexibility. Yet the department-level advantage documented by Perrin (2022a) did not translate into an earlier start for married inventors.

I regress age at first patent on the birth department's share of women in industry, controlling for birth-decade fixed effects and clustering standard errors by birth department (86 clusters; full results in Table 6). The share of women in industry has no detectable effect on the age at first patent in any specification ($p = 0.9669$). Comparing two departments that differ by 9.5 percentage points in female industrial employment (one standard deviation) shifts the predicted patenting age by less than two weeks. Marriage, by contrast, shifts the entire distribution by 7.5 years ($p < 0.001$; R^2 increases from 0.35 to 0.39 when the married indicator enters).

Table 6: OLS regressions of age at first patent on department-level characteristics. Standard errors clustered by birth department. All models include birth-decade fixed effects.

	M1	M2	M3	M4	M5	M6	M7	M8
FemIndLFP	-0.002	0.020	0.003	0.006		0.007	0.015	0.002
	(0.026)	(0.072)	(0.078)	(0.094)		(0.080)	(0.075)	(0.076)
Female literacy		-0.025	-0.020	-0.020	-0.028	-0.058	-0.029	-0.030
		(0.025)	(0.025)	(0.025)	(0.023)	(0.047)	(0.025)	(0.025)
Urban share		-0.043	-0.036	-0.036	-0.035	-0.038	0.014	0.013
		(0.044)	(0.045)	(0.045)	(0.034)	(0.049)	(0.069)	(0.069)
Density		0.000	0.000	0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married			7.517***	7.660***			7.628***	7.656***
			(0.989)	(2.495)			(0.971)	(0.974)
Married × FemIndLFP				-0.004				
				(0.063)				
Wage gap (ind.)					6.390	7.573		
					(8.371)	(9.709)		
Girls school enroll.						4.123		
						(4.936)		
Protestant share							-0.095	-0.082
							(0.088)	(0.087)
Mortality gap							-12.165	-12.137
							(10.453)	(10.044)
Net migration							-47.245	-43.184
							(41.155)	(41.329)
Wage gap (agri.)							8.068	8.372
							(8.674)	(8.303)
Soil quality								0.031
								(0.036)
N	790	780	780	780	780	780	780	780
R^2	0.346	0.345	0.394	0.394	0.346	0.346	0.397	0.398
Birth-decade FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SEs	Dept	Dept	Dept	Dept	Dept	Dept	Dept	Dept

Note: Dependent variable is age at first patent. Standard errors clustered by birth department in parentheses. *** p<0.01, ** p<0.05, * p<0.10. All models include birth-decade fixed effects. Models 5–6 omit the marital status indicator to isolate the department-level channel. Models 7–8 add demographic, cultural, and geographic variables; no department-level variable reaches significance in any specification.

The marriage gap was no smaller in more industrialized departments. I test this by interacting the married indicator with female industrial employment; the interaction coefficient is near zero ($p = 0.9476$). To put a bound on what the data can rule out: comparing a department at the 25th percentile of female employment to one at the 75th percentile, the 95 percent confidence interval on the difference in the marriage gap runs from -1.5 to 1.4 years. The data are precise enough to detect a narrowing of 2.1 years or larger, roughly 27 percent of the 7.5-year gap; any smaller effect would remain undetected. A married woman born in a heavily industrialized department like Nord waited just as long to patent as one born in a rural department like Lozère (Figure C.2 in the appendix confirms this visually: the predicted age gap between married and single women is the same whether female industrial employment is low, medium, or high).

Ten additional department-level variables tell the same story. Female literacy, urbanization, population density, industrial and agricultural wage gaps, girls' school enrollment, Protestant share, mortality gap, net migration, and soil quality all produce coefficients indistinguishable from zero. None reaches significance at the 10 percent level in any specification. Even within individual sectors, sector-specific female employment does not predict when women patent (clothing $p = 0.8295$, textiles $p = 0.8773$, chemistry $p = 0.5041$). Because all inventors in a department share the same employment figures, standard errors could understate the true uncertainty. A wild cluster bootstrap that resamples at the department level confirms these null results (bootstrap $p = 0.7698$).

Marriage shifted the age at first patent by roughly 7.5 years, but no department-level variable narrowed this gap. The gap's stability across departments fits an account in which children consumed time rather than eroded skills (Burnette and Stanfors, 2012): if the delay came from the demands of child-rearing, it should not vary with the structure of local industry. This null has two interpretations. Local economic conditions may genuinely not have affected how long married women waited to patent. But I observe only women who eventually patented, so if local conditions affected whether married women patented at all rather than when, the marriage gap would appear stable across departments even if conditions mattered at the extensive margin.

6.3 Childbearing and patenting

I now examine the relationship between fertility and the timing of patenting, using the 694 ever-married women for whom I observe a birth year and children data from the genealogical sources: 477 mothers with recorded birth dates for at least one child and 208 with no recorded children.[15] These sources capture children registered in the linked genealogical records, which may fall short of a woman's completed fertility when stillbirths or early infant deaths went unrecorded. I cannot test whether the inventors without recorded children differ systematically from those with fertility data, because the genealogical coverage that produces children records is not independent of region or record quality.

6.3.1 Family size and the childbearing window

Most ever-married inventors had small families. About 30.0 percent had no recorded children, 36.9 percent had one or two, and only 33.1 percent had three or more (the full distribution appears in Figure D.6 in the appendix). The mean is 2.0 recorded children; the maximum observed is 9. I exclude the small number of cases with ten or more because these likely reflect data errors or unusually high infant mortality, both of which distort the timing calculations.

Childbearing occupied a well-defined window. Among the 477 mothers with recorded birth dates, the mean age at first birth is 24.6 and the mean age at last birth is 30.6 (Figure 5).[16] The interquartile range for first births spans 21 to 27; for last births, 26 to 35. The average active childbearing span is 6.0 years.[17]

[15] A small number of women have recorded children but lack individual birth-year data for those children. These women appear in the family-size analysis but not in the timing analysis, which is why the mother and no-recorded-children counts do not sum exactly to the total. "No recorded children" does not mean biologically childless: these are women for whom the genealogical sources contain no birth records, whether because they bore no children, because infants died before registration, or because the records did not survive. Comparisons between mothers and women with no recorded children should be read with this caveat.

[16] The full distributions of maternal age at first and last birth appear in Figure D.7 in the appendix.

[17] This average includes mothers of one child, for whom the span is zero by definition. Among mothers of two or more children, the mean span is longer. These figures are consistent with demographic patterns in nineteenth-century

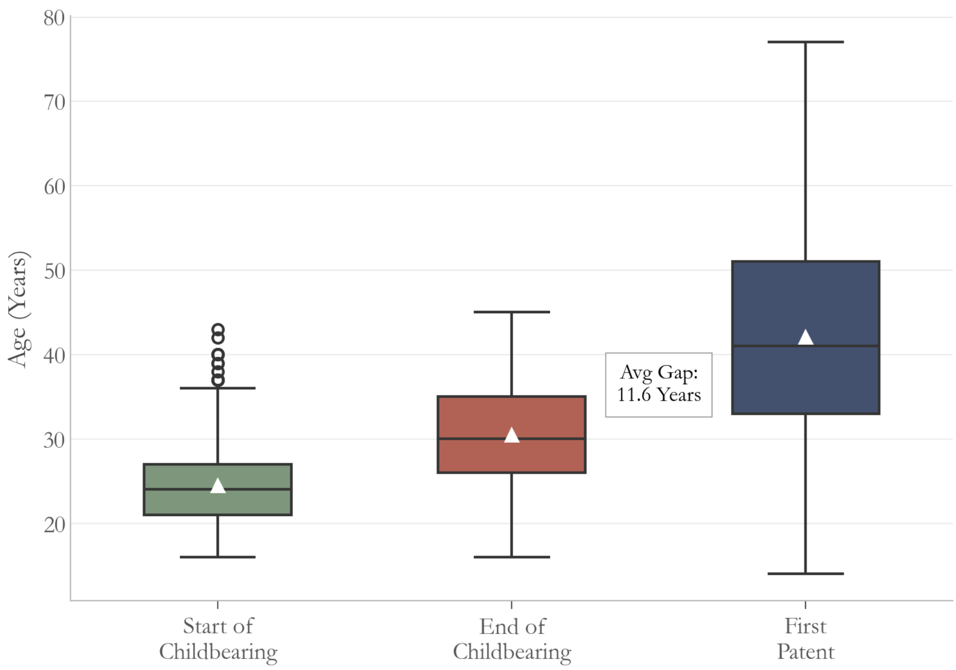


Figure 5: Fertility window and first patent. Maternal age at first birth, last birth, and first patent (N = 477). Invention typically begins after the active childbearing window has closed, with an average gap of 11.6 years between the last birth and the first patent.

6.3.2 The timing of patenting relative to childbearing

For most mothers, invention followed childbearing. Among mothers with recorded birth dates, 94.5 percent filed their first patent after the birth of their first child, and 83.2 percent after the birth of their last child.[18] The average mother gave birth to her last child at age 30.6 and filed her first patent at age 42.2, a mean gap of 11.6 years (median 11; Figure 6 shows the full distribution). Only 13 percent of mothers patented during the active childbearing years, and just 3.6 percent patented before becoming mothers at all.

Women rarely combined active childbearing with patenting (Figure 5). The timing is consistent with the withdrawal-and-return pattern described above. The median mother bore her last child at 30 and filed her first patent at 41, leaving a gap of several years after the most intensive phase of child-rearing had passed. The presence of children under three reduced married women's labor force participation in 1851 France (Grantham, 2012); academic mothers' publications fell during the early-childhood years and recovered as children aged (Kim and Moser, 2025). The 11-year median gap between the last birth and the first patent is consistent with mothers returning to productive work once their youngest children had passed the years of most intensive care. Because genealogical birth records may be incomplete, I use the last observed birth in the linked sources rather than a verified complete fertility history; any under-recording would lengthen the measured gap mechanically.

6.3.3 Family size and the timing of invention

Recorded family size does not clearly predict the age at first patent in this sample. Among the 694 ever-married women with fertility data, the correlation between recorded family size and age at first patent is close to zero ($r = 0.02$, $p = 0.64$; Figure D.9 in the appendix). Women with five or six recorded children filed their first patents at roughly the same age as women with one or two. Recorded family sizes in genealogical records can understate completed fertility, since stillbirths and early infant deaths often went unrecorded; this measurement error could make any true relationship between completed fertility and patenting age harder to detect.

France, where marriage typically occurred in the mid-twenties and childbearing continued into the late thirties for women who had multiple children (Desjardins et al., 1994; Perrin, 2022b).

[18] Figure D.8 in the appendix plots each mother's age at first birth against her age at first patent.

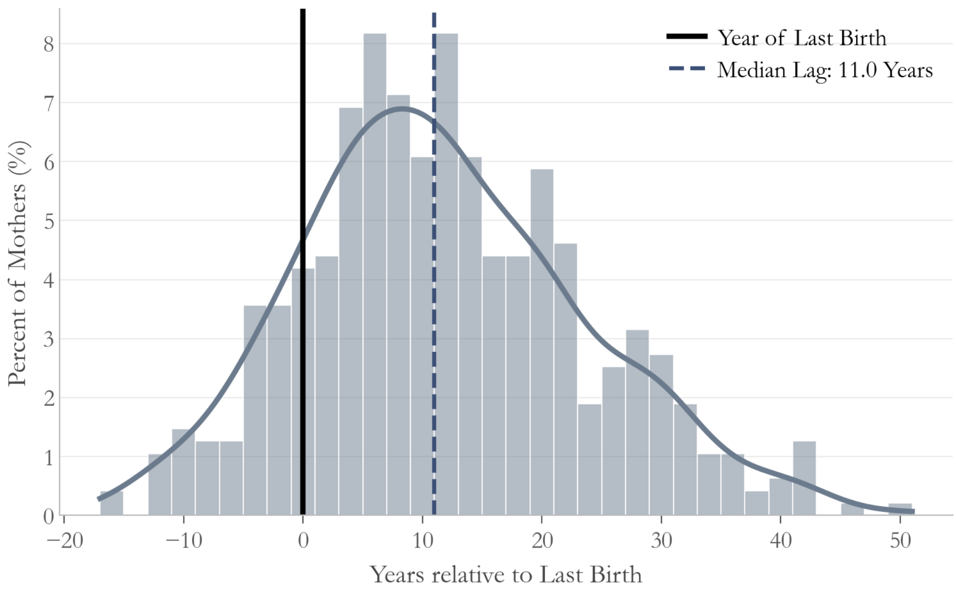


Figure 6: Years from last birth to first patent. Distribution of years elapsed between the last birth and the first patent. The median lag is 11 years, with 83.2% of mothers patenting only after their last child was born.

In other settings, additional children compound the penalty on women's productive lives, both in earnings among Danish women (Kleven et al., 2019) and in publications among academic mothers (Kim and Moser, 2025). The weak gradient observed here could reflect knowledge that did not erode as families grew, but women whom additional children prevented from ever patenting are absent from these records. Family size is also endogenous: women with family-firm support may have had both more children and more resources to sustain inventive activity, so the observed pattern could reflect offsetting forces rather than the absence of a constraint.

Because the observed age at first patent rises little with recorded family size, the gap between the last recorded birth and the first patent mechanically shrinks with additional children: a woman with six recorded children has her last recorded birth at a later age than a woman with two, but both patent at roughly the same observed age, so less time separates the last recorded birth from the patent. The shrinking gap is not a separate finding; it follows from the weak association between recorded family size and age at first patent in this sample.[19] Figure D.10 in the appendix displays the pattern.

First patenting occurred earlier in later birth cohorts. Among mothers born in 1835 or earlier, the mean age at first patent was 48; among those born after, 37. This 11-year gap overstates the true change: later-born mothers cannot be observed at the late ages their predecessors reached, because the analysis window ends in 1900. A woman born in 1860, for example, cannot file a first patent after age 40. If I instead compare only mothers observable over ages 20 to 50, the gap shrinks to about 4 years for the generation that came of age around the 1844 patent reform (n = 274; see Section D.4).

Two broader changes in French society coincided with this shift. The first is the French fertility transition, during which age at first marriage fell across the nineteenth century (Perrin, 2022b) and, in family reconstitutions from four villages in the Haut-Jura, age at last birth fell steadily across cohorts (Desjardins et al., 1994). The second is a sequence of institutional changes: the 1844 patent reform, which spread patent fees across the patent's duration in annual installments (Galvez-Behar, 2019); expanding girls' education in the decades that followed (Perrin, 2022a); and reforms from the 1880s onwards that gave married women progressively greater financial autonomy (Lewis, 1980;

[19]Recorded child mortality does not account for the pattern. The incidence of recorded child loss rises steeply with recorded family size (Figure D.11 in the appendix), but adding the count of children who died before age five to the regression leaves the family-size coefficient little changed; see Figure D.12 in the appendix. This control operates within the recorded sample: births never registered cannot be included.

McMillan, 2000).

In the observed sample, recorded family size was only weakly related to the age at first patent. This specific pattern is not decisive on its own: fertility records may undercount, and I observe only women who eventually patented, not women whom additional children may have kept out of the patent record. The broader pattern of mothers returning to inventive activity after their childbearing years remains consistent with children consuming time rather than eroding the sector-specific knowledge these women had acquired through work. This interpretation cannot rule out that part of the delay reflects the time women needed to secure the resources and, for married women, the legal capacity to file in their own names. Nor did local economic conditions shorten the wait: the gap between the last birth and the first patent was unrelated to female industrial employment ($p = 0.78$). This timing parallels evidence that mothers intensified recorded market work as children entered employment, rather than withdrawing from market work (You, 2020; cf. Tilly and Scott, 1989).

Recorded family size was only weakly related to timing, but it does correlate with sector choice. Among mothers, 30 percent patented in textiles, chemistry, or machinery, compared to 24 percent of women with no recorded children; the pattern reverses in clothing and medicine (20 versus 31 percent; see Section D.3 in the appendix). These differences should not be read as evidence that the local-industry mechanism differs by family status. Interacting the female employment share with a mother indicator in the conditional logit produces a coefficient that is small and statistically indistinguishable from zero ($p = 0.94$). The direction channel operated similarly for mothers and for women with no recorded children.

7 Concluding discussion

I began by asking why women patented in particular sectors and at particular moments of their lives. The evidence in Section 5 and Section 6 gives a two-part answer. What women invented was associated with the gendered structure of local industry. When they patented was associated with marriage, childbearing, and birth cohort.

For sector choice, female employment in the local economy was central. Within departments, the female employment share of local employment was the dominant predictor, roughly twice as strong as total industry size. This result extends the geography-of-innovation literature (Akcigit et al., 2017; Bell et al., 2019). Bell et al. (2019) identify that exposure pattern, but they cannot separate whether the channel runs through knowledge transfer, role-model effects, or access to specialized networks. I provide evidence consistent with the labor market as one such channel: women patented in the sectors that employed women in their birth departments, and the gendered composition of local employment predicted sector choice. The pattern points to occupational knowledge rooted in the industries that employed women locally.

The same local economy, measured at the department level, showed no detectable relationship with when women first patented. Married women patented a median of 12 years later than single women, but no department-level measure of female economic activity narrowed this gap. The parallel with modern evidence is informative: just as generous Danish family policies have not eliminated the contemporary child penalty (Kleven et al., 2019), nineteenth-century industrial structure did not narrow the marriage gap in this sample.

But while the marriage gap was stable across departments, patent timing did shift across cohorts. The generation that came of age around the 1844 patent reform patented about four years earlier than its predecessors. Two broader changes show up in this shift: the French fertility transition (Perrin, 2022b) and the institutional changes that followed the 1844 reform (Galvez-Behar, 2019). Differences in local industry across departments left the marriage gap unchanged, but these broader changes coincide with the cohort shift.

The stability of the marriage gap could also partly reflect the sample restriction to eventual patentees: if childbearing rather than the local economy determined when women could intensify productive work, only those who reached that point appear in the data. These women's combination of intermittent attachment and durable sector-specific knowledge does not fit cleanly into Goldin's (2006) jobs/careers dichotomy. Over a third of women inventors' collaborative ties involved family connections (Merouani, 2026), suggesting that family enterprises were one route through which women acquired sector-specific knowledge despite working intermittently. Regardless of how they first acquired it, the relevant knowledge was tied to the industries that employed women locally and survived the childbearing years.

The fertility evidence sharpens this interpretation. Most mothers patented only after childbearing ended, with a median gap of 11 years between the last recorded birth and the first patent. In the observed sample, recorded family size was only weakly related to the age at first patent, but two limits caution against reading this as a true absence of a child-penalty gradient: genealogical records undercount completed fertility, and I observe only women who eventually patented, not those whom additional children kept out of the patent record altogether. Nineteenth-century artisanal production also differs from the modern academic setting in which a family-size penalty has been documented (Kim and Moser, 2025): knowledge obsolescence at the research frontier, laboratory schedules that demand continuous hours, and a tenure clock that imposes a fixed deadline have no close counterpart in nineteenth-century invention. The pattern I can defend here is narrower than a null contrast: among eventual patentees, the timing of the first patent did not rise steeply with the number of recorded children.

Among women who eventually patented, I find no evidence that craft and artisanal knowledge eroded when they stepped away from production. Artisanal skills were sector-specific rather than firm-specific, so career interruptions did not permanently penalize returning workers (Burnette and Stanfors, 2012). But durability alone did not guarantee a return to invention. Schneider (2026) shows that the mechanization of spinning in Britain destroyed roughly 700,000 part-time jobs held by women and children, whose hand-spinning skills were made redundant by mechanization, producing large-scale and long-term unemployment and underemployment. The French women who patented had both durable knowledge and surviving local industries in which to apply it. French patent law imposed no deadline on when to file, but it did require patentees to exploit their invention once granted, otherwise the patent could be revoked (Galvez-Behar, 2019); a woman would file only when she was both knowledgeable enough to develop a patentable improvement and able to put it into production. Among eventual patentees, first patents clustered after last births, but this pattern cannot identify women kept out of the patent record by permanent withdrawal or by additional childbearing (Section 6).

These findings depend on recovering where women actually lived. 80 percent of patent addresses fall in Paris, but civil records show that only 47 percent of married women married there. Patent agents account for part of the gap: among women who married outside Paris, those with listed agents were far more likely to use Paris addresses. In centralized filing systems, the patent address can reflect the filing channel rather than the inventor's residence. Studies of inventive geography that rely on such records inherit this bias. The linkage itself requires no digitized census microdata, only patent records and a genealogical database covering vital events.

These conclusions rest on a linked sample: only women traceable in genealogical records appear in it. Women with common surnames or poorly preserved local records are underrepresented, and death dates are available for only 53.5 percent of the sample, which limits the lifespan analysis but does not affect the main results on sector choice or patent timing. The conditional logit identifies the female employment share as the dominant predictor of sector choice but cannot isolate the specific pathway through which women acquired sector-specific knowledge. The design also does not reach below the department, so within-department variation across neighborhoods remains outside the comparison. The industrial structure comes from a single 1860 cross-section applied to women born between 1764 and 1881; the results reflect long-lasting regional patterns rather than the specific conditions each woman faced. Undercounting of women's work is a known feature of historical sources (Chiswick and Robinson, 2021); if the 1860 industrial survey disproportionately missed women in female-dominated sectors, the positive coefficient would understate the true relationship. I document robust empirical patterns, not causal effects.

Two extensions would sharpen the results. Linking widows' husbands to death records would test whether legal autonomy, gained when the husband died, rather than the end of childbearing triggered the return to invention; if widows patented soon after bereavement rather than after their last child, the timing mechanism would look more institutional than biological. Linking male inventors to the same local employment data would test whether the direction channel is gendered. If men's sector choice was also predicted by the male employment share of their birth department's industries, the pattern I document is not specifically gendered; it is a general occupational knowledge channel that the gender composition of local industry happens to make visible. If instead total industry size predicted men's sector choice regardless of workforce composition, the gendered structure of production would have shaped women's inventive knowledge in a way it did not shape men's.

To my knowledge, no prior study has linked women inventors to their genealogical records at scale. By building that linkage, I move these women from names in administrative registers to people with places, families, and life courses. I contribute a method as well as a dataset: the method works in any country with comparable patent records and online genealogies. With these data, I show that patent addresses in centralized filing systems cannot be read as residences, a distortion that any map of innovation built from administrative records inherits. I extend the geography of innovation: the female share of local employment, not the total size of local industry, predicted the sectors in which women patented. I connect invention to the child-penalty literature by showing that mothers usually patented after childbearing ended, and that local female economic activity did not narrow the twelve-year marriage gap. Women's invention was local in direction, but life-cycle bound in timing. Practical knowledge linked the two.

References

Akcigit, U., Grigsby, J., and Nicholas, T. 2017. The Rise of American Ingenuity: Innovation and Inventors of the Golden Age: National Bureau of Economic Research w23047

Arrow, K. J. 1962. The Economic Implications of Learning by Doing, *The Review of Economic Studies*, vol. 29, no. 3, 155–73

Baijot, S. and Le Chapelain, C. 2022. Reassessing Women's Participation in Entrepreneurial Activities in the Nineteenth Century: A Review of the Literature, *Œconomia. History, Methodology, Philosophy*, nos. 12–3, 405–42

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J. 2016. *The Lifecycle of Inventors*, http://www.ssrn.com/abstract=2838018 (date last accessed 11 November 2025)

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J. 2019. Who Becomes an Inventor in America? The Importance of Exposure to Innovation*, *The Quarterly Journal of Economics*, vol. 134, no. 2, 647–713

Berger, T. and Prawitz, E. 2024. Inventors among the 'Impoverished Sophisticate', *The Journal of Economic History*, vol. 84, no. 4, 1175–1207

Bonneuil, N., Bringé, A., and Rosental, P.-A. 2008. Familial components of first migrations after marriage in nineteenth-century France, *Social History*, vol. 33, no. 1, 36–59

Boter, C. 2021. The Emergence of the Dutch Housewife Revised. How Shifts in Local Labour Market Structures Shaped Dutch Unmarried Women's Labour Force Participation, 1812–1929, *Historical Life Course Studies*, vol. 10, 130–34

Boter, C. and Woltjer, P. 2020. The impact of sectoral shifts on Dutch unmarried women's labor force participation, 1812–1929, *European Review of Economic History*, vol. 24, no. 4, 783–817

Braudel, F. 1984. *The Perspective of the World*, Civilization and Capitalism, 15th–18th Century, London, Phoenix Press

Burnette, J. 2011. *Gender, work and wages in industrial revolution Britain*, Cambridge studies in economic history, Cambridge, Cambridge University Press

Burnette, J. and Stanfors, M. 2012. Was there a family gap in late nineteenth century manufacturing? Evidence from Sweden, *The History of the Family*, vol. 17, no. 1, 31–50

Chanteux, A. 2023. 'Les filles d'Athéna : Femmes et brevets au XIXe siècle en France', These de doctorat, Paris, EHESS

Chanteux, A. 2024. Images of Female Inventors: Women and Patents in Nineteenth-Century France, *Artefact. Techniques, histoire et sciences humaines*, no. 21, 397–427

Chanut, J.-M., Heffer, J., Mairesse, J., and Postel-Vinay, G. 2000. *L'industrie française au milieu du XIXe siècle: Les enquêtes de la Statistique générale de la France*, Paris, Éditions de l'EHESS

Charpentier, A. and Gallic, E. 2018. *Internal Migrations in France in the Nineteenth Century*, https://arxiv.org/abs/1807.08991

Charpentier, A. and Gallic, E. 2020. Can historical demography benefit from the collaborative data of genealogy websites?, *Population*, vol. 75, no. 2, 379–408

Chiswick, B. R. and Robinson, R. H. 2021. Women at work in the United States since 1860: An analysis of unreported family workers, *Explorations in Economic History*, vol. 82, 101406

Craig, B. 2016. *Female Enterprise Behind the Discursive Veil in Nineteenth-Century Northern France*, London, Springer

Craig, B. 2020. A Constant Presence: The Businesswomen of Paris, 1810–1880, pp. 113–36, in Aston, J. and Bishop, C. (eds.), *Female Entrepreneurs in the Long Nineteenth Century: A Global Perspective*, Cham, Springer International Publishing

Desjardins, B., Bideau, A., and Brunet, G. 1994. Age of mother at last birth in two historical populations, *Journal of Biosocial Science*, vol. 26, no. 4, 509–16

Dupin, C. 1826. *Carte figurative de l'instruction populaire de la France*, Bruxelles

Galvez-Behar, G. 2019. The patent system during the French industrial revolution: Institutional change and economic effects, *Jahrbuch für Wirtschaftsgeschichte/Economic History Yearbook*, vol. 60, no. 1, 31–56

Goldin, C. 2006. The Quiet Revolution That Transformed Women's Employment, Education, and Family, *American Economic Review*, vol. 96, no. 2, 1–21

Grantham, G. 2012. Occupational, Marital, and Life-Cycle Determinants of Women's Labor Force Participation in Mid Nineteenth-Century Rural France, *Feminist Economics*, vol. 18, no. 4, 97–119

Harvey, D. A. 2000. Mothers and Breadwinners: Gender and Working Class Identity in Alsace, 1821-1936, *Left History: An Interdisciplinary Journal of Historical Inquiry and Debate*, vol. 7, no. 2, 7–27

Humphries, J. and Schneider, B. 2019. Spinning the industrial revolution, *The Economic History Review*, vol. 72, no. 1, 126–55

Humphries, J. and Weisdorf, J. 2015. The Wages of Women in England, 1260–1850, *The Journal of Economic History*, vol. 75, no. 2, 405–47

Jung, T. and Ejermo, O. 2014. Demographic patterns and trends in patenting: Gender, age, and education of inventors, *Technological Forecasting and Social Change*, vol. 86, 110–24

Kesztenbaum, L. 2008. Places of Life Events as Bequestable Wealth: Family Territory and Migration in France, 19th and 20th Centuries, pp. 155–84, in Bengtsson, T. and Mineau, G. P. (eds.), *Kinship and Demographic Behavior in the Past*, Dordrecht, Springer Netherlands

Khan, B. Z. 2016. Invisible Women: Entrepreneurship, Innovation, and Family Firms in Nineteenth-Century France, *The Journal of Economic History*, vol. 76, no. 1, 163–95

Khan, B. Z. 2020. *Inventing Ideas: Patents, Prizes, and the Knowledge Economy*, Oxford, New York, Oxford University Press

Kim, S. and Moser, P. 2025. Women in Science. Lessons From the Baby Boom, *Econometrica*, vol. 93, no. 5, 1521–60

Kleven, H., Landais, C., and Søgaard, J. E. 2019. Children and Gender Inequality: Evidence from Denmark, *American Economic Journal: Applied Economics*, vol. 11, no. 4, 181–209

Landes, D. S. 1969. *The unbound Prometheus: Technological change and industrial development in Western Europe from 1750 to the present*, Cambridge, Cambridge University Press

Lewis, H. D. 1980. The Legal Status of Women in Nineteenth-Century France, *Journal of European Studies*, vol. 10, no. 39, 178–88

McBride, T. M. 1977. The Modernization of "Woman's Work", *The Journal of Modern History*, vol. 49, no. 2, 231–45

McMillan, J. F. 2000. *France and Women, 1789–1914: Gender, Society and Politics*, London, Routledge

Merouani, Y. 2026. *What Happens When Inventors Collaborate? Evidence from 19th-Century Patents*

Merouani, Y. and Perrin, F. 2022. Gender and the long-run development process. A survey of the literature, *European Review of Economic History*, vol. 26, no. 4, 612–41

Merouani, Y. and Perrin, F. 2024. Myths and Biases: An Exploration of Women's Historical Patenting Activities, pp. 253–76, in Le Chapelain, C. (ed.), *Nineteenth Century Businesswomen*, Cham, Springer Nature Switzerland

Merouani, Y. and Perrin, F. 2026a. *Pioneering Women: The Determinants of Women's Inventive Productivity*

Merouani, Y. and Perrin, F. 2026b. Women Inventors: On the Origins of the Gender Patenting Gap, *Journal of Economic History*, Accepted for publishing 2026

Mokyr, J. 2005. *The gifts of Athena: Historical origins of the knowledge economy*, Princeton N.J, Princeton University Press

Nuvolari, A., Tortorici, G., and Vasta, M. 2023. British-French Technology Transfer from the Revolution to Louis Philippe (1791–1844): Evidence from Patent Data, *The Journal of Economic History*, vol. 83, no. 3, 833–73

Nuvolari, A. and Vasta, M. 2017. The geography of innovation in Italy, 1861–1913: Evidence from patent data, *European Review of Economic History*, vol. 21, no. 3, 326–56

O'Brien, P. and Keyder, Ç. 1978. *Economic Growth in Britain and France 1780 - 1914: Two Paths to the Twentieth Century*, London, Allen & Unwin

Perrin, F. 2022a. Can the historical gender gap index deepen our understanding of economic development?, *Journal of Demographic Economics*, vol. 88, no. 3, 379–417

Perrin, F. 2022b. On the origins of the demographic transition: Rethinking the European marriage pattern, *Cliometrica*, vol. 16, no. 3, 431–75

Riboud, M. 1985. An Analysis of Women's Labor Force Participation in France: Cross-Section Estimates and Time-Series Evidence, *Journal of Labor Economics*, vol. 3, no. 1, 177–200

Sarada, S., Andrews, M. J., and Ziebarth, N. L. 2019. Changes in the demographics of American inventors, 1870–1940, *Explorations in Economic History*, vol. 74, 101275

Schneider, B. 2026. Technological Unemployment in the British Industrial Revolution: The Destruction of Hand-Spinning*, *Past & Present*, vol. 270, no. 1, 105–49

Stanfors, M., Leunig, T., Eriksson, B., and Karlsson, T. 2014. Gender, productivity, and the nature of work and pay: Evidence from the late nineteenth-century tobacco industry, *The Economic History Review*, vol. 67, no. 1, 48–65

Tilly, L. A. and Scott, J. W. 1989. *Women, Work and Family*, New York, Routledge

Whittington, K. B. 2011. Mothers of Invention?: Gender, Motherhood, and New Dimensions of Productivity in the Science Profession, *Work and Occupations*, vol. 38, no. 3, 417–56

You, X. 2020. Women's labour force participation in nineteenth-century England and Wales: Evidence from the 1881 census enumerators' books, *The Economic History Review*, vol. 73, no. 1, 106–33

Zijdeman, R. L., van Leeuwen, M. H. D., Rébaudo, D., and Pélissier, J.-P. 2014. Working women in France, nineteenth and twentieth centuries. Where, when, and which women were in work at marriage?, *The History of the Family*, vol. 19, no. 4, 537–63

A Linking algorithm and sample validation

This appendix provides the full technical specification of the linking pipeline summarized in Section 3.

A.1 Data collection and filtering

A.1.1 Building searches and parsing results

For married women with maiden names, each search URL encodes the following parameters:

- **Maiden name** (surname field): the inventor's birth surname, which should appear in birth and marriage records.
- **Given name** (first name field): allows for variant spellings and compound names common in the period.
- **Spouse surname** (spouse surname field): the marital surname, which filters results to women married to men with that name.
- **No spouse given name**: I do not specify the husband's given name, because this is rarely recorded in patent documents. The male individual sharing the surname could also be a son or other family member.
- **Variant tolerance**: the search interface permits minor spelling variations. I set this to no variation, except to allow for additional given names.

For single women, I search using only the birth surname and given name, with no spouse fields. For married women without a maiden name, I search using the marital surname in the surname field and the inventor's given name.

I save the HTML of each search results page locally and parse it to extract structured data. Each result row contains a mix of visible text and HTML comments encoding metadata. I extract:

- **Individual and spouse names**: the name of the person in the record and any spouse name.
- **Event years**: birth, marriage, and death years, drawn from both explicit year fields and date strings.
- **Event locations**: places associated with birth, marriage, death, and the record itself.
- **Source metadata**: the type of archival source (civil vital records, census, military, cemetery, notarial, user-submitted family trees).
- **Relationship context**: whether the record describes the person directly, their spouse, a parent, or a cited individual.

Some records provide only partial information; a marriage record may contain a marriage year and place but no birth or death data. This variation in completeness is central to the challenge of reconstructing biographies.

A.1.2 Initial filtering rules

I eliminate records that are clearly implausible on hard criteria. I reject records whose given name has no overlap with the inventor's given name, because any plausible match must share at least one name in common. I also reject records where the maiden surname similarity falls below 60 percent; this threshold allows for spelling variants while excluding plainly different surnames. If a birth year implies the inventor was younger than 14 or older than 85 when filing her first patent, I reject the record. If a death year precedes the first patent by more than three years, or the last patent by more than two years, I reject the record, because death dates drawn from cemetery registers or family trees may carry recording errors of one to two years and because heirs occasionally filed patents on behalf of a recently deceased inventor. When both birth and marriage years are present and derive from high-confidence sources, I reject records implying marriage before age 14 or after age 90.

These rules bind only when dates come from reliable sources such as civil vital records, military records, or notarial acts. When dates derive from family trees or other user-submitted sources, I flag the record as suspicious but allow it to proceed to the scoring stage.

A.2 Record linking: scoring, clustering, and selection

A.2.1 Record-level scoring

For each surviving record, I compute a quality score with five components, each ranging from near-zero to 1.0.

The first component measures **name match quality**. I compare the maiden surname and given name in the record to those in the patent using a combination of Levenshtein ratio for surnames and token-set overlap for given names. A perfect match on both receives 1.0; partial matches receive intermediate scores between 0.2 and 1.0.

The second component measures **spouse match quality**. I compare the spouse surname in the record to the marital surname from the patent. A perfect match scores 1.0, and if the spouse's given name also matches, the evidence is stronger. Partial matches score between 0.2 and 1.0.

The third component measures **place consistency**. I compare the department and country in the patent's address field to places mentioned in the record, including birth, marriage, death, and record-keeping locations. Because patent addresses are unreliable (inventors often provide temporary addresses in Paris or use patent agents' addresses), I weight place match less heavily than name and spouse match.

The fourth component measures **source reliability**. I classify each record's archival source into a reliability ladder: civil vital records receive weight 1.0, census and military records 0.7, notarial records 0.6, cemetery registers 0.5, user-submitted family trees 0.45, and unknown sources 0.5.

The fifth component measures **temporal plausibility**. I evaluate whether event timing makes sense relative to patent dates. A marriage year more than two years after the first patent is penalized, since the patent indicates the woman was already married. Birth and death years must fall within reasonable age-at-patenting bounds.

I combine the five components into a single row weight using a multiplicative model:

$$w_{\text{row}} = \exp\left(\sum_i \alpha_i \ln(\max(\epsilon, s_i))\right)$$

where s_i are the five component scores, α_i are exponents controlling the relative importance of each dimension, and $\epsilon = 10^{-6}$ prevents taking the logarithm of zero. The multiplicative form ensures that a record must score reasonably well on all dimensions to receive meaningful weight. The α_i values and all other algorithm parameters are documented in the replication code.

A.2.2 Micro-clustering

Records about the same person may appear in multiple archival sources with slight variations in spelling, dates, or places. I group them into clusters representing distinct biographical hypotheses.

I use a two-level blocking strategy. First, I partition records by normalized spouse surname (removing accents and capitalization). Within each surname block, I further partition by a signature of the spouse's given name, capturing the first given name and, if multiple names are present, the initial of the second. I merge signatures within each block if they are sufficiently similar, using a token-set ratio threshold of 85 percent.

Within each block, I perform temporal micro-clustering. Two records are compatible if their effective birth, marriage, and death years either agree within tolerance (plus or minus one year for birth and death, plus or minus two years for marriage) or at least one is missing. I enforce strict agreement only when both records provide high- or medium-confidence years for the same event.

I use a union-find algorithm to identify connected components of compatible records. Each component becomes a candidate cluster. Clusters with full date information (strong anchors) take precedence over those with only contextual information (weak anchors). Context-only clusters survive only if no anchored clusters exist for the same inventor; otherwise, I attempt to attach their records to nearby anchored clusters based on place and spouse-name similarity.

A.2.3 Consensus-building and cluster selection

For each cluster, I compute a consensus biography through weighted hierarchical voting. For birth, marriage, and death years, I first vote among high- and medium-confidence years from civil vital records, military records, and similar reliable sources. If none exist, I fall back to low-confidence years from family trees. For death years, cemetery records serve as a proxy when civil death records are unavailable. I use a tolerance-aware mode calculation: two years are considered the same if they fall within plus or minus one year for birth and death, or plus or minus two years for marriage. The winning year is the one with the highest aggregate weight within its tolerance window. I record both the consensus year and the share of total weight supporting it.

For places, I restrict the vote to records supporting the winning event year, ensuring that the consensus birth place comes from records that agree on the birth year. If no place is available from year-consistent records, I fall back to all records in the cluster. Place strings carrying structured geographic information (commune, postal code, department, and country) receive three times the voting weight, because they provide more precise and verifiable location data.

I then apply cluster-level vetoes: age at first patent must fall between 14 and 85, death must not precede the first patent by more than three years or the last patent by more than two, and age at marriage must fall between 14 and 90. A cluster producing an implausible consensus is discarded entirely.

For each inventor, I rank surviving clusters by total weight (the sum of row weights after deduplication by archival link). I select the highest-weighted cluster, with two caveats. First, if the gap between the top two clusters is less than 15 percent, I flag the result as ambiguous; if the second cluster has higher-confidence years, I may select it instead. Second, if the consensus birth or death year has low internal agreement (share below 50 percent), I flag the result as medium quality rather than high quality.

A.3 Algorithm adaptations

A.3.1 Single women

Single women appear in patent records with only a birth surname and given name. I adapt the four-stage architecture by replacing the spouse signal with parent names as the primary genealogical anchor. The parent match component receives the highest weight of any dimension (1.2, compared to 1.0 for source reliability and name match). Parent names serve the same disambiguation role as spouse names: a record naming the same father and mother as other records in the cluster is far more likely to describe the same person, because parent-name coincidences across unrelated individuals are rare.

In the clustering stage, I block records by normalized birth surname and sub-block by parent-name signatures rather than spouse-name signatures. The similarity threshold adapts to context: when one parent's surname matches at 95 percent or higher across two records, I lower the threshold for the other parent from 90 to 80 percent, because the strong match on one parent provides confidence that a slight discrepancy on the other reflects transcription variation rather than a different family.

Because the single algorithm lacks the independent spouse signal, I impose a stricter quality gate. A link starts as unconfirmed and must earn promotion through at least one independent signal: the inventor's name is rare enough that the search returned 50 or fewer results, the patent address matches a non-Parisian genealogical location, the winning cluster contains records from official civil registers, the cluster draws on five or more parent-linked records with a combined score above 2.0, or a co-inventor's given name matches a family member in the cluster. Only promoted links are retained, all classified as high quality. There is no medium or ambiguous tier.

A.3.2 Married women without maiden names

Searching genealogical databases with only a marital surname creates a circularity problem: the search returns women married to men with that surname, so the spouse-surname field trivially matches the search term. The maiden name, which normally provides the independent check, is unavailable.

The key insight is that while the spouse surname is circular, the spouse's given name is not. If a search for women married to someone named "Dupont" returns thirty records, and twenty-one name the husband "Jean," that consistency is strong evidence those twenty-one records describe one real woman. I require that at least 70 percent of records in the winning cluster agree on the spouse's given name.

Because this algorithm has the weakest identifying signal, I apply the strictest acceptance thresholds. The maximum age at patent drops from 85 to 75. The minimum cluster weight rises from 1.0 to 1.5. Every winning cluster must contain at least three records after deduplication, must hold at least 70 percent of the total weight across all clusters for that inventor, and must achieve at least 80 percent internal consensus on its birth or death year. I require that the consensus include either a birth year or a death year; clusters with only a marriage year are rejected, because without a temporal anchor the match cannot be validated against patent dates. Searches returning more than 200 results after initial filtering are skipped entirely. All surviving links are classified as high quality.

A.4 Validation sampling design

I assess reliability through manual verification of a stratified sample of 60 inventors covering both married and single women, designed to oversample difficult cases. Simple random sampling would miss rare but informative failure modes, so I use a stratified design with known selection probabilities.

I first construct a risk score for each inventor combining eight signals of matching difficulty: the algorithm's cluster score, name and spouse match quality, source reliability, missing dates, death year confidence, number of competing clusters, reliance on user-contributed family trees, and surname commonness. Higher scores indicate cases where the algorithm might struggle. I divide inventors into four groups: edge cases (extremely common surnames, cemetery-only death records, inventors missing both birth and death dates) and three risk-based groups among the remainder (ambiguous: top 25 percent of risk; high: 40th to 75th percentile; medium: bottom 40 percent). I draw the sample in two rounds of 30 inventors each, with more drawn from riskier strata in both rounds. The second round extends coverage to single inventors linked through the maiden-name algorithm, because these cases lack the spousal anchor that helps disambiguate married matches.

Within each stratum, I sample with probability proportional to risk, so an inventor with a risk score of 0.8 is twice as likely to be selected as one with a score of 0.4. This ensures difficult cases appear in the sample while maintaining known selection probabilities for every inventor.

For each sampled inventor, I independently verify birth, marriage, and death years against original document scans in departmental archives, genealogical platforms, and original patent filings at the INPI, reading the handwritten records directly rather than relying on transcribed metadata, and without reference to the algorithm's output. Patent documents prove especially useful for confirming identity, because the original filings record the inventor's marital status ("mademoiselle," "veuve," "madame née"), domicile address, and co-inventors, all of which I cross-check against the genealogical match. I compare the algorithm's dates to this gold standard using a tolerance of plus or minus two years for all three events, a standard accommodation in historical demography for cemetery records that may record burial rather than death dates and for variation in how individuals report ages across archival sources.

I report 95 percent Clopper-Pearson exact binomial confidence intervals for each event type, because these intervals make no distributional assumptions and remain valid at small sample sizes.[20]

[20] With 60 observations the intervals are informative: the lower bound for birth-year accuracy exceeds 86 percent.

A.5 Linked-sample representativeness

This section provides the distributional comparisons that support the balance tests reported in the main text (Table 1 and Table 2). I show the sectoral and occupational composition of each linked sample alongside its reference population, and I disaggregate by marital status to check whether the patterns differ across single, married, and widowed women.

A.5.1 Sectoral distribution across linked and reference groups

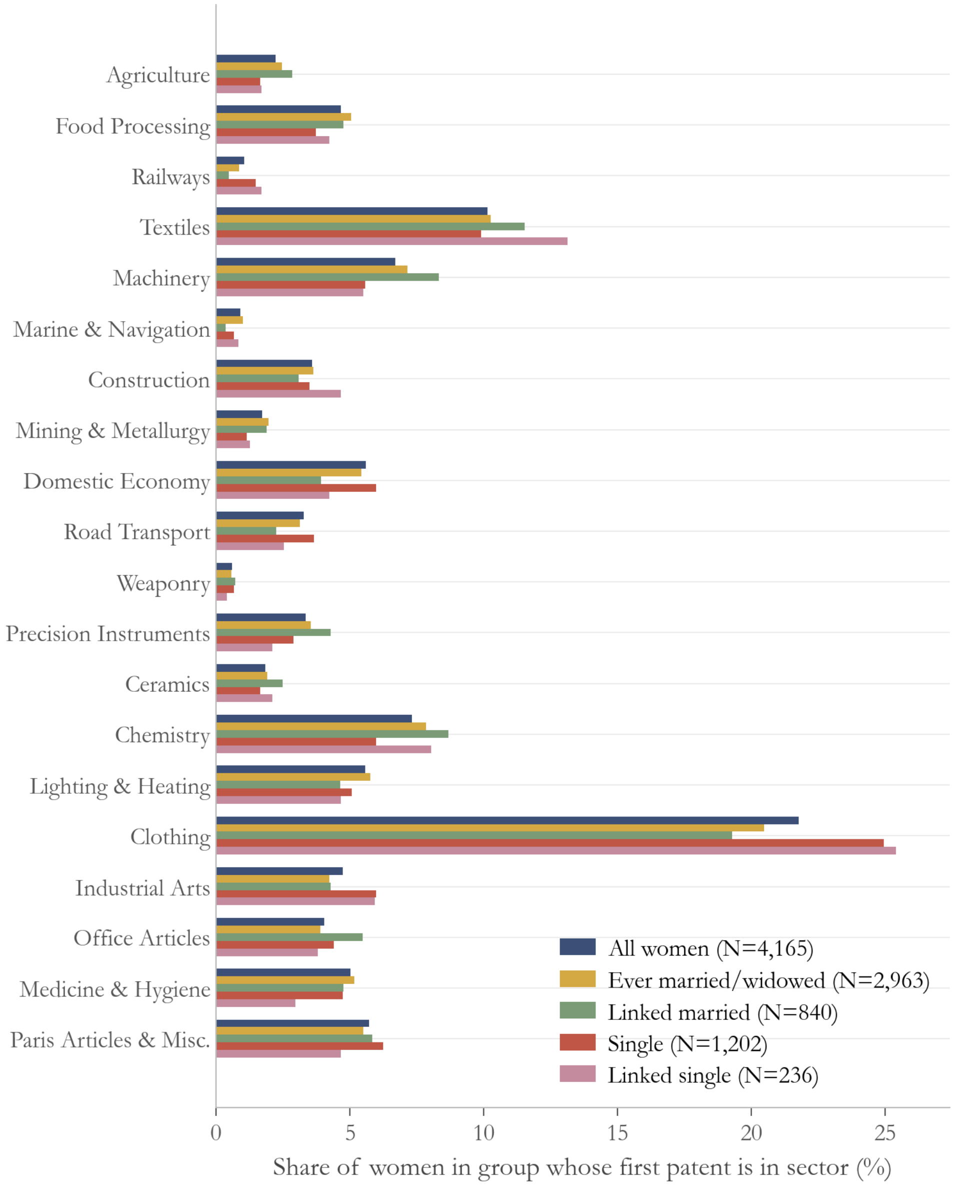


Figure A.1: Sector distribution of women inventors' first patents for five groups: all women in the sample, women who were ever married or widowed, linked married or widowed women, single women, and linked single women. Values show, for each group, the percentage of women whose first recorded patent falls in each technological sector.

Figure A.1 shows the sector distribution of first patents across all 4,165 women in the sample, the 2,963 ever married or widowed, the 840 linked ever-married, the 1,202 single, and the 236 linked single. The dominant sectors look similar across groups. Clothing accounts for 19 to 25 percent everywhere, with the variation running along marital-status lines rather than the linked-versus-unlinked margin: 20.5 percent of the ever married or widowed and 19.3 percent of the linked ever-married, 25.0 percent of the single and 25.4 percent of the linked single. Textiles ranges from

10 to 13 percent across groups. Chemistry runs from about 6 to 9 percent. The sectors where the linked samples depart most visibly from their reference populations, Office Articles and Domestic Economy among married women, Textiles and Medicine and Hygiene among single women, are discussed in the main text alongside the formal balance tests in Table 1 and Table 2.

A.5.2 Sectoral distribution by marital status

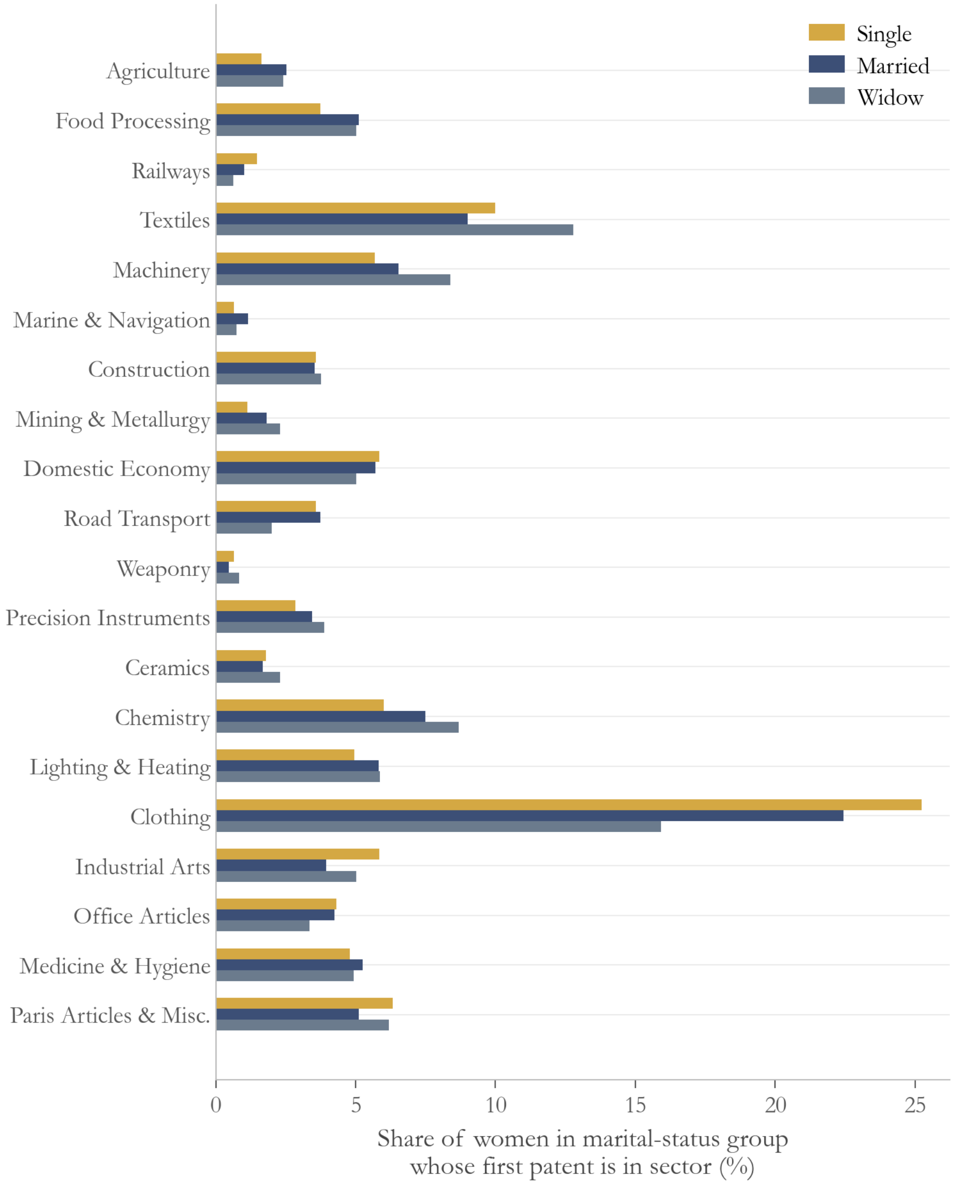
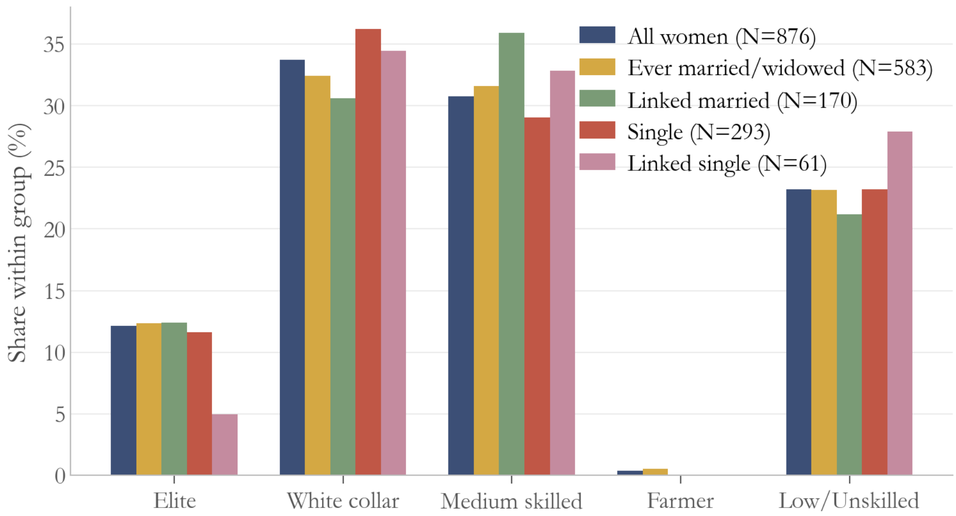


Figure A.2: Sector distribution of women inventors' first patents within marital-status groups. For each marital-status category (Single, Married, Widow), the bars represent the share of women in that category whose first recorded patent lies in each sector; percentages within each marital-status group sum to 100.

Figure A.2 breaks the full inventor population out by marital status at the time of the first patent. Single women show the highest concentration in Clothing, at about 25 percent, compared to 22 percent for married women and 16 percent for widows. Widows are more heavily represented in Textiles (about 13 percent versus 10 percent for single women and 9 percent for married women), Chemistry (about 9 percent versus 6 percent for single women), and Machinery (about 8 percent versus 6 percent). These patterns may reflect changing economic circumstances after a husband's death, but the differences do not dominate the overall distribution: Clothing remains the largest sector for all three groups.

A.5.3 Skill distribution (HISCLASS)

Panel A: Skill distribution (observed) by group



Panel B: Skill distribution (observed) within marital-status groups

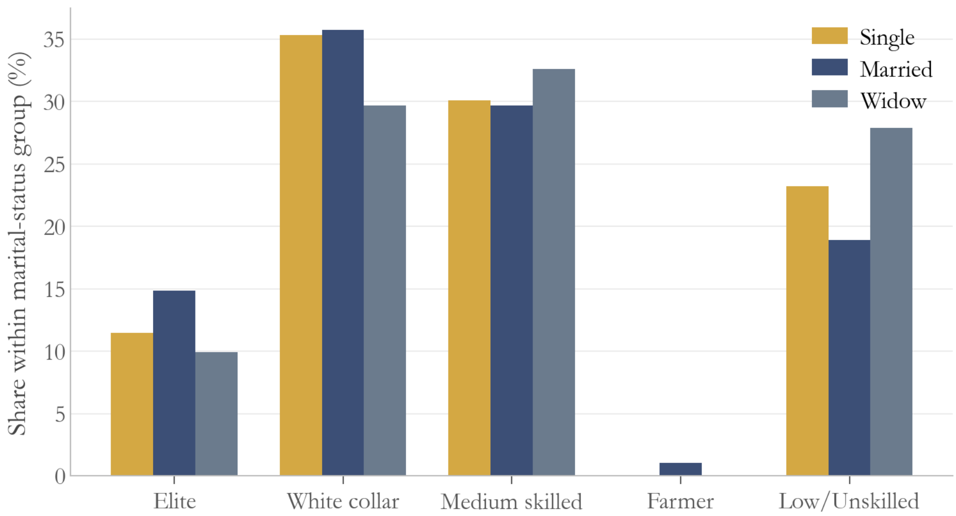


Figure A.3: Occupation-based profiles of women inventors (observed skill level). Panel A compares the skill distribution across five groups: all women inventors with occupation data (N = 876), women ever married or widowed (N = 583), linked ever-married women (N = 170), single women (N = 293), and linked single women (N = 61). Panel B shows how skill distribution varies by marital status (single, married, widow) within the full sample. Both panels display only women for whom occupation information was recorded in patent applications, which represents a minority of all inventors.

Figure A.3 reports the HISCLASS skill distribution of women inventors. Panel A covers the five comparison groups used throughout this paper. White collar workers make up the largest category in nearly every group, ranging from about 31 percent among linked ever-married women to about 36 percent among single women. Medium skilled workers come second, at roughly 29 to 36 percent. Low or unskilled workers account for 21 to 28 percent. The linked ever-married sample shifts modestly from white collar toward medium skilled work relative to the broader married and widowed population, a pattern consistent with women in occupations with more common names being harder to link. The linked single sample broadly mirrors the full single population in the two largest categories but shows a sharp drop in Elite representation, from about 12 to 5 percent. With only 61 observations carrying occupation data, this estimate is imprecise.

Panel B disaggregates the full population by marital status. Single and married women look similar, with white collar workers at about 35 to 36 percent. Widows stand apart: the white collar share falls to about 30 percent and the low or unskilled share rises to about 28 percent, compared to about 19 percent for married women. This likely reflects changes in economic position after a husband's death.

A.5.4 Major HISCO distribution

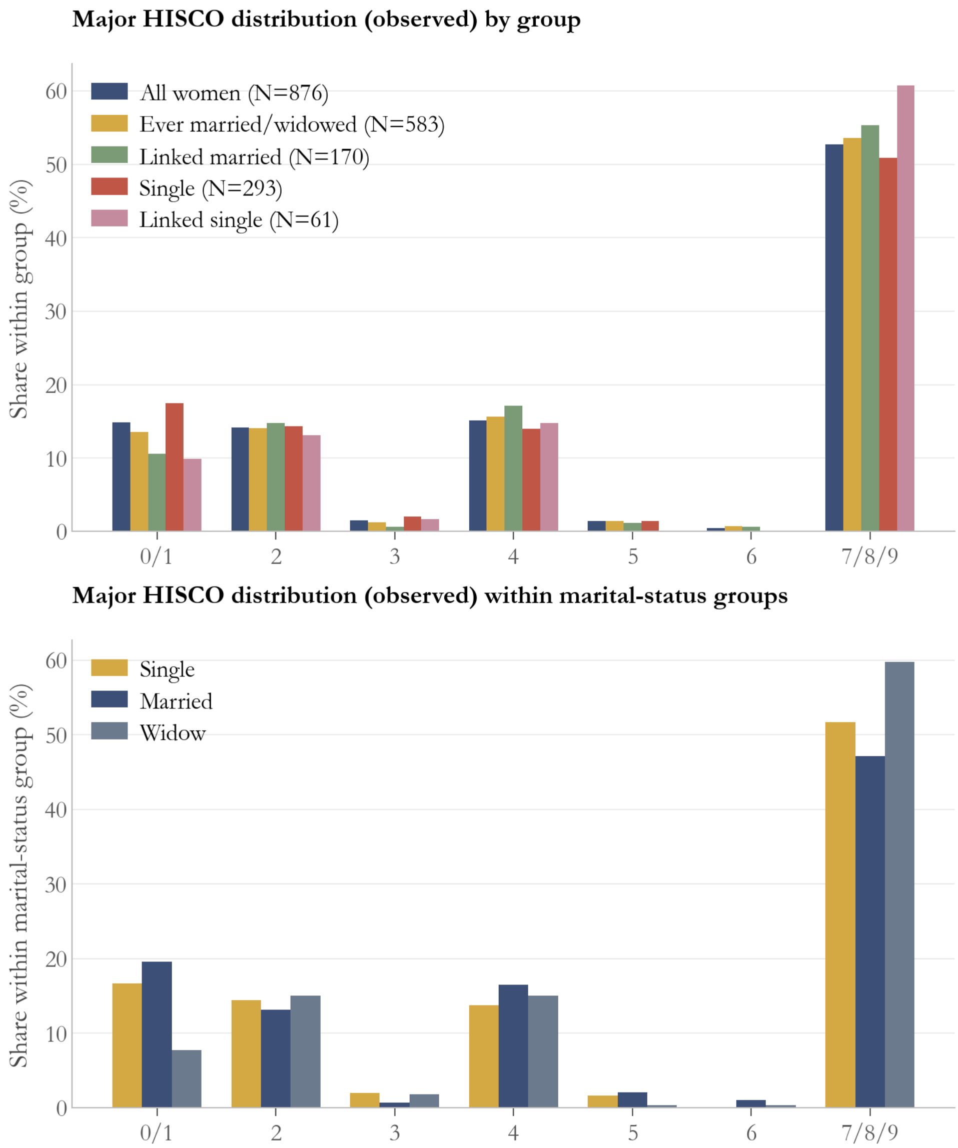


Figure A.4: Occupation-based profiles of women inventors (observed Major HISCO). Panel A compares the distribution across five groups: all women inventors with occupation data (N = 876), women ever married or widowed (N = 583), linked ever-married women (N = 170), single women (N = 293), and linked single women (N = 61). Panel B shows the distribution within marital-status groups for the full sample.

Figure A.4 reports the Major HISCO distribution across the same five groups and by marital status. Production and related workers (Major HISCO 7/8/9) dominate everywhere, accounting for roughly 47 to 61 percent of women with recorded occupations. Professional and technical workers (0/1) come second in most groups, at about 10 to 17 percent. The linked ever-married sample has a somewhat lower professional and technical share (about 11 percent) than the ever married or widowed population (about 14 percent) and a correspondingly higher production share (about 55 versus 54 percent). The linked single sample shows the largest shift: production workers rise to about 61 percent and professional and technical workers fall to about 10 percent. As with the

HISCLASS comparisons in Figure A.3, the small sample sizes counsel against reading too much into these differences.

B Geography

B.1 Patent address reliability: detailed evidence

This section provides detailed evidence supporting the address-reliability findings reported in Section 4.

Table B.1 reports the cross-tabulation of marriage locations and patent addresses. Among the 383 inventors who married outside Paris, 64.5 percent have patent addresses in Paris. Among the 344 inventors who married in Paris, 96.8 percent have Paris patent addresses. Patent addresses are systematically concentrated in Paris regardless of where women married.

Table B.1: Marriage location vs. patent address

Marriage Location	Patent in Paris	Patent Outside Paris	Total
Paris	333 (96.8%)	11 (3.2%)	344
Outside Paris	247 (64.5%)	136 (35.5%)	383
Total	580	147	727

Notes: Sample restricted to 727 married inventors with valid marriage and patent location data. Percentages shown are row percentages. Paris is defined as department 75 (Seine).

Table 3 and Figure B.1 reinforce this pattern. Parisian Residents (both in Paris) show a median distance of 0.0 kilometres, as expected for residential addresses. Non-Paris Residents (both outside Paris) show a median of 1.6 kilometres and a mean of 66.6 kilometres, with 69.9 percent of inventors within 50 kilometres of their marriage location. These distances are consistent with addresses that reflect residence.

Paris-Address Non-Parisians (married outside Paris, patent in Paris) tell a different story. The median distance is 212.7 kilometres and the mean is 247.9 kilometres. Only 6.1 percent of inventors filed patents within 50 kilometres of where they married, and 56.3 percent filed from locations more than 200 kilometres away. Department-level match rates show a similar pattern: 100.0 percent of Parisian Residents match at the department level, 71.3 percent of Non-Paris Residents match, but 0 percent of Paris-Address Non-Parisians match by definition, since their marriage and patent departments differ and patents are concentrated in Paris.

Table B.2 refines this classification by incorporating death locations to distinguish genuine migrants from administrative users. These patterns indicate that patent addresses do not simply reflect residential locations. Not all of the 64.5 percent of inventors married outside Paris who list Paris patent addresses migrated to the capital. The question is which of the 247 Paris-Address Non-Parisians genuinely migrated to Paris and which used Paris addresses for administrative convenience.

Death records provide one way to distinguish genuine migrants from administrative users of Paris addresses. Among the 127 Paris-Address Non-Parisians for whom I observe death locations, roughly half died in Paris (68) and roughly half died outside Paris (59). Those who died in Paris likely migrated to the capital at some point after marriage. Those who died outside Paris show no evidence of permanent migration, yet they still filed patents with Paris addresses.

I label these subgroups Confirmed Migrants and Administrative Users. Neither label is definitive: dying in Paris does not prove residence at the time of patenting, and dying outside Paris does not rule out temporary residence in the capital. But the Administrative Users are the more informative group. These 59 women married outside Paris, listed Paris addresses on their patents, and died outside Paris. The median distance from marriage location to patent address is 265.3 kilometres, and the median time from marriage to patent filing is 20 years. The profile is consistent with long-term residents of other departments who used Paris addresses for administrative purposes rather than genuine migrants.

Patent agents explain part of this pattern. Among inventors married outside Paris, 77.6 percent of those with listed patent agents have Paris addresses, compared to 50.8 percent of those without agents (difference: 26.7 percentage points, $p < 0.001$). Even among inventors without listed patent agents, half used Paris addresses despite marrying elsewhere, pointing to widespread use of Paris-based intermediaries, including lawyers, family members, and business contacts, beyond officially recorded patent agents.

Table B.2: Refined classification of inventors by Paris connection

Group	N	% of Sample	Interpretation
Parisian Residents	333	45.8%	Married, patented, and (when observed) died in Paris
Confirmed Migrants	68	9.4%	Married outside Paris, patented in Paris, died in Paris
Administrative Users	59	8.1%	Married outside Paris, patented in Paris, died outside Paris
Unknown Trajectory	120	16.5%	Married outside Paris, patented in Paris, death unobserved
Non-Paris Residents	136	18.7%	Married and patented outside Paris
Other	11	1.5%	Other combinations or missing data
Total	**727**	**100.0%**	

Notes: Death location observed for 388 inventors (53% of sample). Administrative Users labeled as such because temporary Paris residence during patenting cannot be ruled out.

Administrative Users represent a lower bound on non-residential patent addresses: at least 8.1 percent of inventors used Paris addresses without residing there permanently, under the assumption that death location reflects long-term residence. If the 120 inventors in the Unknown Trajectory group resemble the observed Paris-Address Non-Parisians, approximately 56 additional inventors used Paris addresses without permanently relocating, yielding an upper bound of 15.8 percent. This extrapolation assumes that inventors with unobserved deaths resemble those with observed deaths, which may not hold if Paris deaths are differentially recorded.

The 136 Non-Paris Residents (both married and patented outside Paris) provide a useful benchmark, though the group is not homogeneous. Among the 97 who married and patented in the same department, the median distance is 0.8 kilometres and the mean is 8.4 kilometres, consistent with addresses that reflect residence. The remaining 39 married in one department and patented in another, with a median distance of 128.9 kilometres, consistent with the minority of long-distance internal moves documented in the demographic literature. Outside the Paris administrative orbit, patent addresses track residential locations for the majority of inventors.

The contrast with Paris-Address Non-Parisians is sharp. Non-Paris Residents who stayed within their department show a median distance of 0.8 kilometres; Paris-Address Non-Parisians show 212.7 kilometres. This difference is consistent with the Paris patent system, rather than residential location alone, driving patent address patterns.

The 236 single inventors in the linked sample cannot enter this marriage-based test, but birth location provides a weaker anchor, because birth precedes patenting by several decades and allows substantial life-course migration. Even so, the same pattern appears: 78 percent have Paris patent addresses while only 13 percent were born in Paris, and the median distance from birth to patent address is 249.0 kilometres (N = 203). The divergence between civil records and patent addresses in this independent group corroborates the pattern observed among married women.

Three limitations bear on these estimates. I observe death locations for only 53 percent of the sample, and deaths in Paris may be better documented than deaths outside the capital, which would bias the Administrative Users count downward. I cannot distinguish temporary Parisian residence during patent filing from purely administrative use of Paris addresses. And marriage location is an imperfect proxy for residence at the time of patenting, because the median 19-year gap from marriage to patent for Paris-Address Non-Parisians allows substantial life-course migration.[21]

[21] Quantitative studies using genealogical, civil-registration, and linked microdata suggest that nineteenth-century French families were typically clustered in relatively small areas, with a minority undertaking long-distance moves. Using collaborative genealogies, Charpentier and Gallic (2018) find that for descendants of individuals born 1800–1804, about 62% of children are born in exactly the same commune as their ancestor and a further 19% within 20 km, leaving roughly 18% as long-distance migrants. Studies of the 3,000 Families Survey show similar patterns: among 5,627 women, only 55 are observed moving between vital events, and of these 55 female migrants 30 move less than 12 km; for men, 486 of 888 migrants move less than 12 km and 701 less than 50 km (Bonneuil et al., 2008). Kesztenbaum (2008), using life-event locations and military registers, shows that half of families have a mean distance to the family barycentre below 4 km and classifies family "territories" with low dispersion (0–4 km) and medium dispersion (4–20 km), indicating that many life events occur within a narrow radius. Long-distance migration to cities was nonetheless important, particularly for women entering urban domestic service: McBride (1977) argues that domestic service was "the most important spur, outside of marriage, to feminine urban migration," and documents substantial long-range flows of young women from rural Brittany, Normandy, Alpine regions and other departments into Paris and Lyon.

Despite these limitations, the evidence consistently points in one direction: patent addresses systematically over-represent Paris. Comparing Paris death rates (62 percent) to Paris patent addresses (80 percent) suggests a bias on the order of 17 percentage points. The institutional explanation is straightforward. Under the nineteenth-century French patent laws, applications were filed with a central office in Paris and could be submitted by the inventor or by a *mandataire* (agent or representative). Galvez-Behar (2019) notes that many applicants from outside Paris used Paris-based agents whose addresses appeared in the records. This centralized filing, which Section 4 contrasts with the decentralized Italian system, likely produced the systematic address patterns I observe.

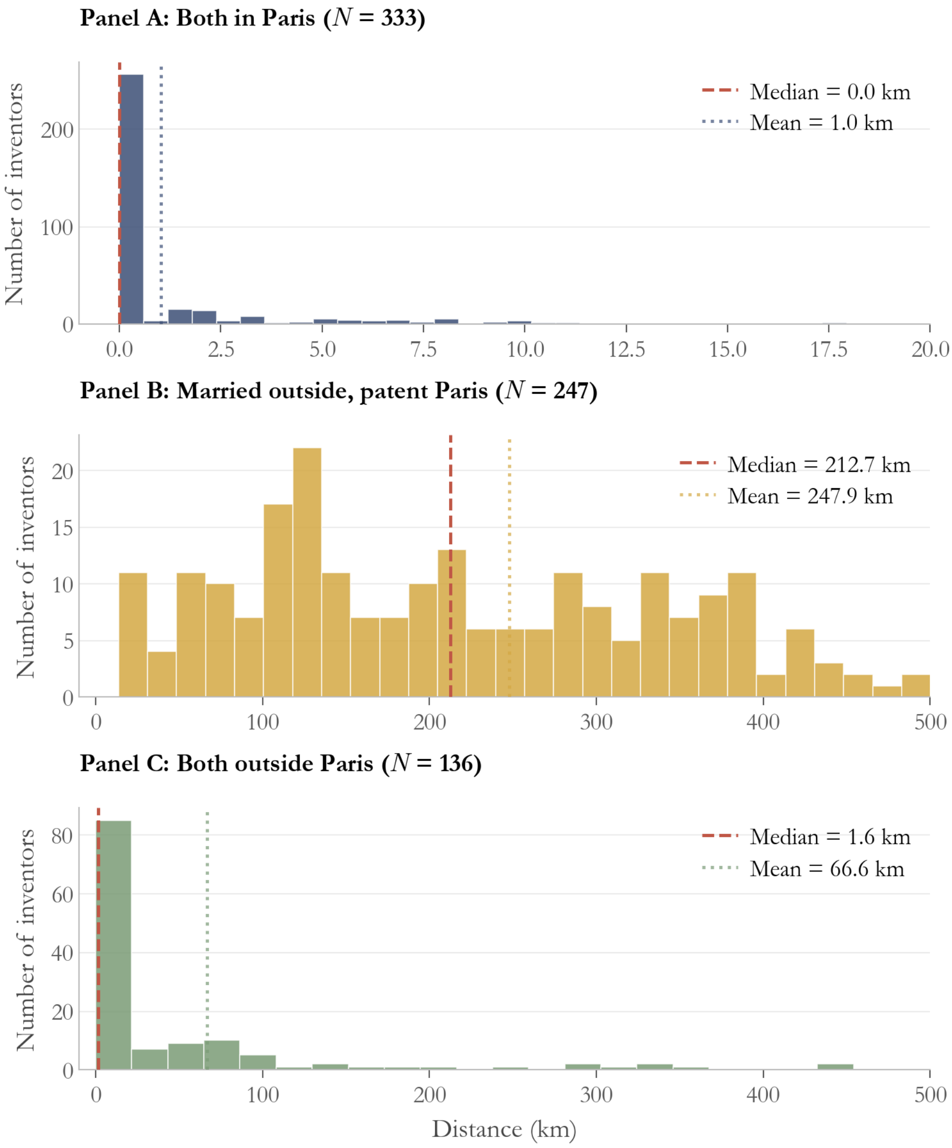


Figure B.1: How far is a patent address from where the inventor married? Geodesic distances (km) between each inventor's marriage location and her patent address, stratified by whether marriage and patent fell inside or outside Paris (department 75, Seine). Panel A: both events in Paris (median 0.0 km). Panel B: married outside Paris but patent address in Paris (median 212.7 km). Panel C: both events outside Paris (median 1.6 km). If patent addresses reflect residence, all three panels should cluster near zero. The wide dispersion in Panel B suggests that many Paris patent addresses do not correspond to the inventor's place of residence. Sample restricted to married inventors with valid coordinates for both locations.

These studies together suggest that short-distance moves and sedentariness were common, while long-distance urban migration characterized a selected minority, such as servants, rather than the median woman.

B.2 Migration pathways

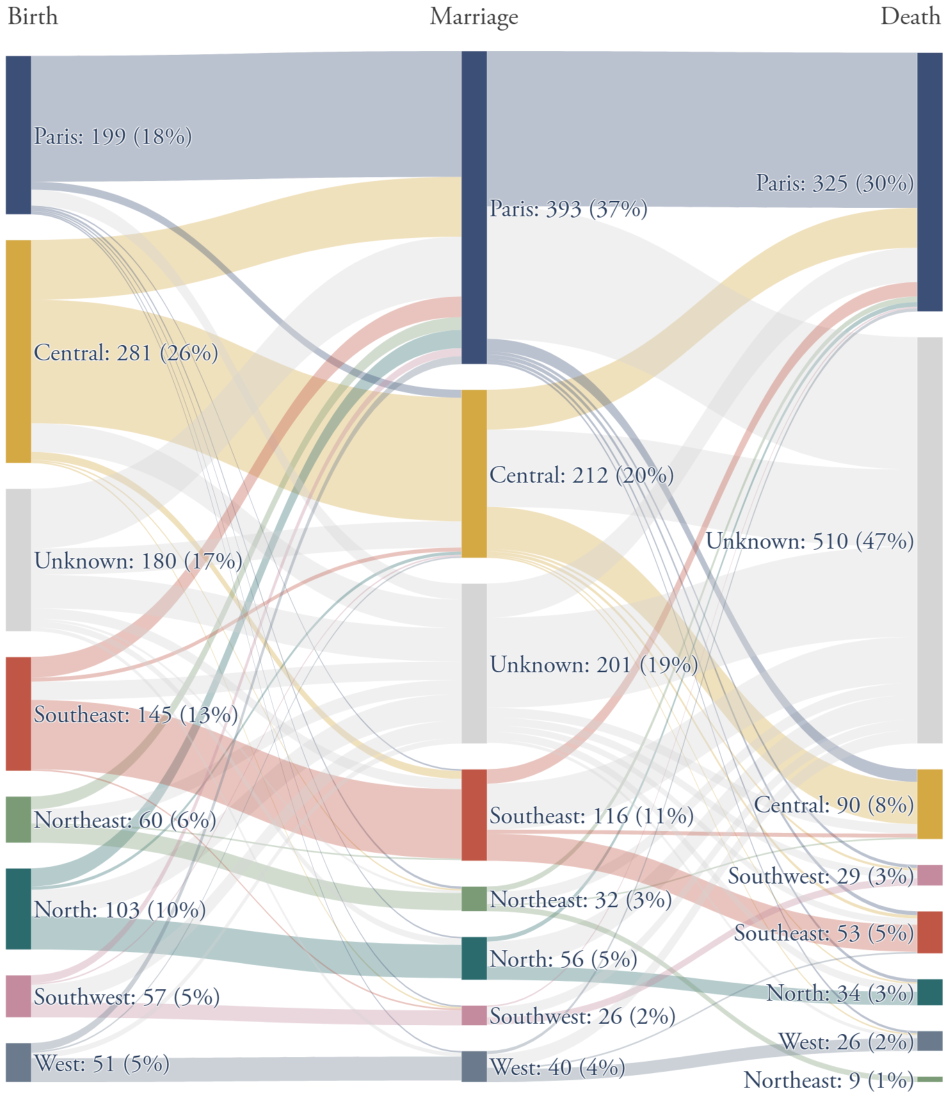


Figure B.2: Geographic flows for 1,076 linked French women inventors, 1791–1900. Flow widths are proportional to the number of women on each path. Of the 201 women classified as Unknown at marriage, 135 are single with no marriage record; the remaining 66 are married or widowed whose marriage location could not be geocoded. The large Unknown share at death (47 percent) reflects the scarcity of death records in genealogical sources. Paris is a 0.3-degree radius around the city center; other regions partition metropolitan France.

B.3 The North-South divide

Section 4 documented the North-East/South-West development gradient first mapped by Dupin (1826) and refined by Braudel (1984). Nuvolari et al. (2023) use the Braudel line (Rouen to Geneva) to characterize the spatial concentration of French patenting between 1791 and 1844. I test whether this divide predicts women's inventor rates.

I classify each department as North or South of two lines: the Dupin line (St. Malo to Geneva) and the Braudel line (Rouen to Geneva), based on the latitude of the department's inventor-weighted centroid relative to the line equation. The two classifications agree for 90 percent of departments.

Table B.3: Women inventor rates and the North-South divide. Dependent variable is inventors per 100,000 female residents (1861). Paris excluded.

	Dependent variable: Inventors per 100,000 Women				
	(1) Baseline	(2) + Braudel	(3) + Dupin	(4) Braudel only	(5) Dupin only
Female literacy rate	0.042**	0.027	0.023		
	(0.014)	(0.015)	(0.020)		
Share women in agriculture	-0.048	-0.041	-0.044		
	(0.029)	(0.034)	(0.035)		
North of Braudel line		1.176		2.770***	
		(1.009)		(0.478)	
North of Dupin line			1.111		2.639***
			(1.304)		(0.465)
Observations	81	81	81	84	84
R^2	0.344	0.366	0.358	0.279	0.293
Note:					*p<0.05; **p<0.01; ***p<0.001

Note: Dependent variable is women inventors per 100,000 female residents (1861). Robust standard errors (HC1) in parentheses. The Braudel line runs from Rouen to Geneva; the Dupin line from St. Malo to Geneva. Departments are classified as North or South based on the latitude of their inventor-weighted centroid relative to the line. Paris (Seine) excluded from all models.

Both lines are strongly significant when entered alone (Braudel $p < 0.001$, $R^2 = 0.28$; Dupin $p < 0.001$, $R^2 = 0.29$), confirming that the North-South divide predicts inventor rates. When department-level literacy and agricultural employment are added as controls, both lines become insignificant (Braudel $p = 0.24$; Dupin $p = 0.39$), suggesting that these continuous measures already capture most of the North-South gradient in inventor production.

C Local industry and invention: supplementary materials

C.1 Patent technology class to 1860 industry concordance

Table C.1: Patent technology class to 1860 industry concordance.

Patent Sector	N	Tier	1860 Industry	Rationale
Food Processing	37	1	Food processing	Direct match
Textiles	97	1	Textiles	Direct match
Construction	26	1	Building	Direct match
Mining & Metallurgy	13	1	Mining + Metallurgy	Direct match
Ceramics	16	1	Ceramics	Direct match
Chemistry	72	1	Chemicals	Direct match
Lighting & Heating	37	1	Lighting	Direct match
Clothing	155	1	Clothing	Direct match
Agriculture	27	2	Agriculture	Direct match
Railways	7	2	Transport	Railway transport
Marine & Navigation	4	2	Transport	Maritime transport
Road Transport	18	2	Transport	Road transport
Machinery	58	3	Metallurgy	Metal-working machinery
Domestic Economy	33	3	Furnishings	Household furnishings
Precision Instruments	31	3	Science & art	Precision instruments/science
Industrial Arts	40	3	Science & art	Artistic/scientific instruments
Office Articles	43	3	Science & art	Education/scientific equipment
Medicine & Hygiene	33	3	Chemicals	Pharmaceuticals/chemicals
Paris Articles & Misc.	38	3	Luxury goods	Luxury goods (Paris specialties)

Note: Tier 1 = direct semantic match between patent sector and 1860 census industry. Tier 2 = clean indirect mapping (agriculture, transport). Tier 3 = judgment call, validated by minor patent class data. N = number of inventors in birth-department sample (total = 785). "11. Weaponry" excluded (N<10).

Table~C.1 maps each of the twenty patent technology classes to the most closely corresponding 1860 industrial survey category. The 1860 *Enquête Industrielle* records sixteen manufacturing sectors under their original French names. Throughout this paper I use English translations: *Textile* (Textiles), *Extractive* (Mining), *Métallurgie* (Metallurgy), *Construction* (Construction), *Cuir* (Leather), *Bois* (Wood), *Céramique* (Ceramics), *Chimique* (Chemicals), *Bâtiment* (Building), *Éclairage* (Lighting), *Ameublement* (Furnishings), *Habillement* (Clothing), *Alimentation* (Food processing), *Transport* (Transport), *Science et art* (Science & art), *Luxe* (Luxury goods).

Direct mappings are unambiguous: patent class Clothing maps to survey industry Clothing, Textiles to Textiles, and so on. Three additional technology classes map directly but appear under a different variable format in the survey: agriculture is recorded separately from the sixteen manufacturing sectors, and transport aggregates three patent categories.

The remaining seven technology classes involve judgment calls guided by patent minor-class data. For these, I use the minor patent classification (98 minor classes) to assign individual inventors to the most appropriate 1860 industry rather than forcing a single mapping for the entire major class. For example, a sewing-machine patent (minor class 05.7) is assigned to Textiles rather than Metallurgy, and a furniture patent (minor class 09.4) is assigned to Furnishings rather than a generic category. Funeral objects (minor class 19.5, ten inventors) have no clear 1860 counterpart and are excluded from the minor-class concordance variant.

C.2 Permutation test results

Table C.2: Permutation tests for Location Quotient match scores. 10,000 permutations. One-sided empirical p-values.

	Tier 1			All Tiers		
	Observed	z	p	Observed	z	p
Female LQ	1.7076	2.639	0.008***	1.6122	-0.288	0.575
Male LQ	1.7425	2.576	0.009***	1.9001	-0.158	0.545
FemShare	0.6399	-0.972	0.833			
Residual	2.417	1.529	0.064*			
Non-Paris	1.4608	1.794	0.044**			
Paris only	2.4751	1.498	0.178			

Note: 10,000 permutations. One-sided p-values. *** p<0.01, ** p<0.05, * p<0.10. FemShare and Residual use female employment share and OLS residuals (female on male employment) instead of LQ.

Table C.2 reports all permutation test variants. The core test uses the female Location Quotient for inventors in the eight directly mapped technology classes (10,000 permutations). Additional rows show the male LQ (which also rejects the null, because female and male employment are correlated at $r = 0.97$ across departments), the female employment share (FemShare = female/(female+male), which does not reject the null because the permutation test cannot separate the gendered component using across-department variation), the OLS residual (female employment regressed on male, then residuals used as match scores; marginally significant at $p = 0.064$), and the Paris split. The Paris-only result is not significant ($p = 0.178$), consistent with Paris having diversified industry across all sectors.

C.3 Conditional logit: cluster bootstrap and leave-one-out

Standard errors in the conditional logit are based on individual likelihood contributions. Because inventors from the same department share identical employment rates, the effective number of independent observations is closer to the 86 departments with inventors than the 453 individuals. I implement a department-level cluster bootstrap with 999 replications, resampling departments with replacement. Table C.3 compares the default and bootstrapped standard errors for Model (2) (female and male employment).

Table C.3: Cluster bootstrap standard errors for Model (2). Confidence intervals and p-values from bootstrap percentile method (999 replications).

Coefficient	Default SE	Bootstrap SE	Inflation	95% CI	p
β_{fem} (+0.0722)	0.007	0.029	4.3×	[0.0445, 0.1684]	0.0000
β_{male} (-0.0526)	0.007	0.029	4.2×	[-0.1485, -0.0259]	0.0000

The standard errors inflate by a factor of approximately five. Both coefficients remain significant: the bootstrap percentile 95% confidence intervals exclude zero for both the female employment coefficient ([0.0445, 0.1684]) and the male employment coefficient ([-0.1485, -0.0259]). The coefficient was negative in all 999 replications. The two coefficients are statistically distinguishable from each other (bootstrapped Wald $\chi^2 = 4.6$, $p = 0.0322$). The main text focuses on the preferred Model (3) decomposition, which is robust to clustering in every variant (Table 5). All robustness rows in Table 5 use the same department-level cluster bootstrap. The Wald test for Model (2), which asks whether β_{fem} and β_{male} are distinguishable, is significant in the baseline ($p = 0.0322$) and in the larger robustness variants, but loses power in smaller subsamples where only 191 to 343 inventors are spread across fewer departments. The preferred female employment share coefficient, which captures the gendered composition of local industry rather than the absolute magnitudes that the Wald test contrasts, remains significant in every variant.

A leave-one-out exercise drops each industry from the choice set and re-estimates Model (2). Table C.4 reports the results.

Table C.4: Leave-one-out exercise for Model (2). Each row drops one industry from the choice set and re-estimates.

Dropped industry	β_{fem}	Δ from baseline	N inventors
None (baseline)	0.072	—	453
Food processing	0.077	+0.005	416
Building	0.074	+0.002	427
Ceramics	0.072	-0.000	437
Chemicals	0.077	+0.004	381
Lighting	0.076	+0.004	416
Mining & metallurgy	0.070	-0.002	440
Clothing	0.171	+0.098	298
Textiles	0.080	+0.008	356

Dropping any sector except Clothing produces β_{fem} estimates between 0.069 and 0.080. Dropping Clothing (the largest sector, 298 of 453 inventors) increases the coefficient to 0.171. This increase means that the remaining sectors display an even stronger gendered pattern; Clothing moderates the aggregate estimate because its large mass dilutes the cross-sectoral variation. The sign and significance are stable regardless of which sector is excluded.

C.4 IIA and mixed logit

The independence of irrelevant alternatives (IIA) requires that the ratio of choice probabilities between any two sectors is unaffected by the inclusion or exclusion of other sectors. A Hausman-McFadden test rejects this assumption for 4 of 7 testable sectors at the 5 percent level: Clothing, Ceramics, Chemicals, and Food processing. I estimate a mixed logit that draws each inventor's female employment share coefficient from a normal distribution and estimates the mean and standard deviation across the sample via simulated maximum likelihood, relaxing IIA without requiring the researcher to impose a nesting structure.

Under the mixed logit (N = 440), the mean female employment share coefficient is 2.33 (SE = 0.28, $p < 0.001$), larger than the conditional logit estimate of 1.82. The standard deviation across inventors is 3.13, confirming meaningful heterogeneity in responsiveness to local female employment. The mean effect is unambiguously positive.

C.5 Temporal mismatch

I test for temporal mismatch by interacting the female employment share with the absolute distance between each inventor's approximate year of knowledge acquisition (birth year plus twenty) and 1860. The interaction coefficient is negative ($\beta = -0.03$, $p = 0.037$): the female employment share effect does weaken for cohorts born far from 1860. The main effect remains large ($\beta = 2.36$, $p < 0.001$), and the net effect at the sample's mean distance of 18 years is 1.81, close to the baseline estimate of 1.82.

Splitting the sample confirms this pattern. Inventors whose exposure falls within twenty years of 1860 (born 1820 to 1860, N = 284) produce a female employment share coefficient of 2.01; those twenty to forty years away (N = 139) produce 1.89. Both are highly significant and similar in magnitude. 61 percent of the analysis sample falls within twenty years of 1860, and the result holds across both cohorts that are large enough to estimate precisely.

C.6 Technology-class heterogeneity

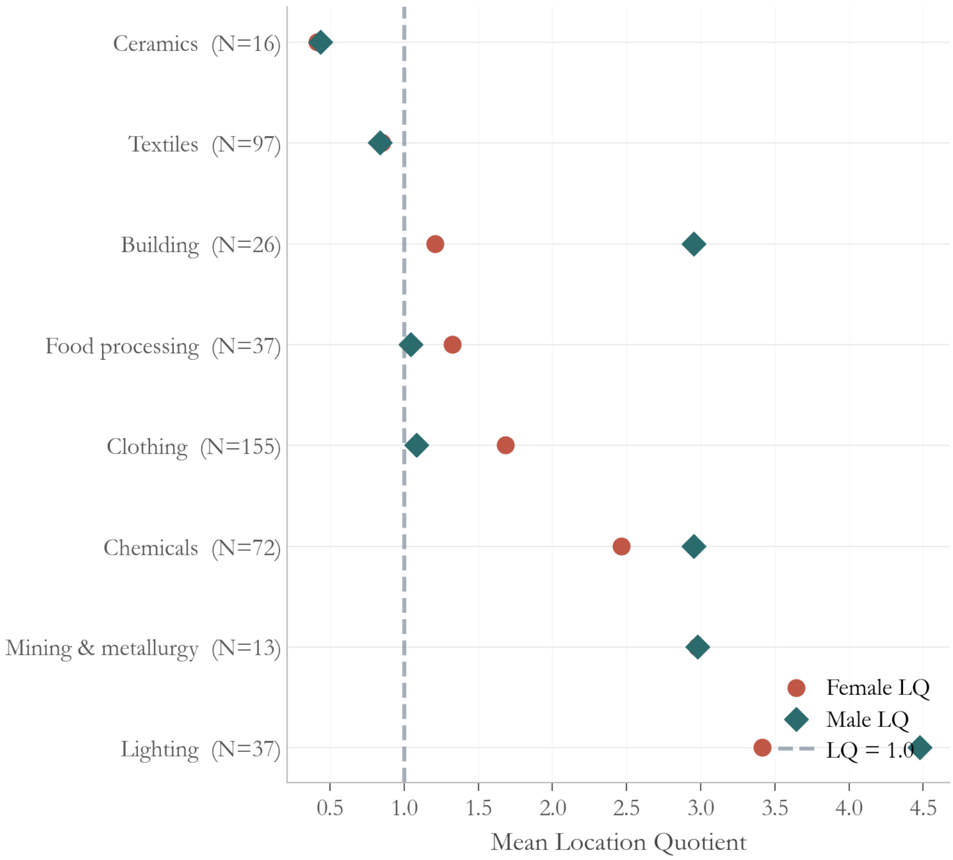


Figure C.1: Mean Location Quotient by technology class for the eight directly mapped industries. Blue circles: female LQ. Orange diamonds: male LQ. The dashed line marks LQ = 1 (no specialization).

Lighting (N = 37) shows the strongest geographical concentration (mean female LQ = 3.41). Chemicals (N = 72, LQ = 2.47) and Clothing (N = 155, LQ = 1.68) also exceed one. Textiles (N = 96, LQ = 0.85) and Ceramics (N = 16, LQ = 0.41) fall below one: their inventors do not come from departments where these sectors are especially concentrated. The textile result reflects the fact that textile employment was so widespread nationally that no department was truly specialized in it. The conditional logit, which exploits within-department variation in gender composition rather than across-department specialization, still identifies textile as female-dominated relative to other sectors.

C.7 Movers: detailed results

Among the 671 inventors with both birth and marriage departments, 245 are movers and 426 are stayers (birth department equals marriage department). I classify movers into three groups: 134 moved to Paris (born outside, married in department 75), 88 moved between two non-Paris departments, and 23 moved from Paris to a department outside the capital.

The conditional logit horse race includes employment shares from both the birth and marriage departments as competing predictors. For all movers with directly mapped sectors, the marriage-department coefficient is $\beta = 0.042$ ($p < 0.001$) and the birth-department coefficient is $\beta = 0.003$ ($p = 0.63$), with a Wald test rejecting equality ($p = 0.0002$). For the 46 non-Paris movers alone, the marriage coefficient weakens to $\beta = 0.021$ ($p = 0.091$) and the Wald test fails to reject equality ($p = 0.617$). The small non-Paris subsample limits statistical power: with 46 inventors and twelve alternatives, the effective variation for identifying the birth-versus-marriage distinction is limited.

C.8 Timing regressions

Table 6 reports the eight timing regressions. Models 1 through 4 progressively add controls. The share of women in industry has no effect in any specification (smallest $p = 0.78$). The married indicator adds 7.5 years to the predicted age ($p < 0.001$) and increases R^2 from 0.35 to 0.39; all department-level variables add no explanatory power. Models 5 and 6 replace or supplement female industrial employment with the gender wage gap and girls' school enrollment, neither of which reaches significance. Models 5 and 6 omit the married indicator to isolate the department-level channel; the lower R^2 relative to Models 3 and 4 reflects this exclusion, not any explanatory power of the department variables. Models 7 and 8 extend the baseline (Model 3) with demographic, cultural, and geographic variables: Protestant share, mortality gap, net migration, agricultural wage gap, and soil quality. None reaches significance. The R^2 rises from 0.39 to 0.398, confirming that these additional variables add no explanatory power beyond the marital status indicator and birth-decade fixed effects.

Sector-specific timing regressions confirm the aggregate null. Among the 155 clothing inventors, the habillement employment rate in the birth department is unrelated to age at first patent ($p = 0.83$). The same holds for textile inventors ($p = 0.88$) and chemistry inventors ($p = 0.50$). A wild cluster bootstrap using Rademacher weights (999 replications) returns $p = 0.77$ for the female industrial employment coefficient in Model 3, confirming that the null is not an artifact of the asymptotic clustered standard error approximation.

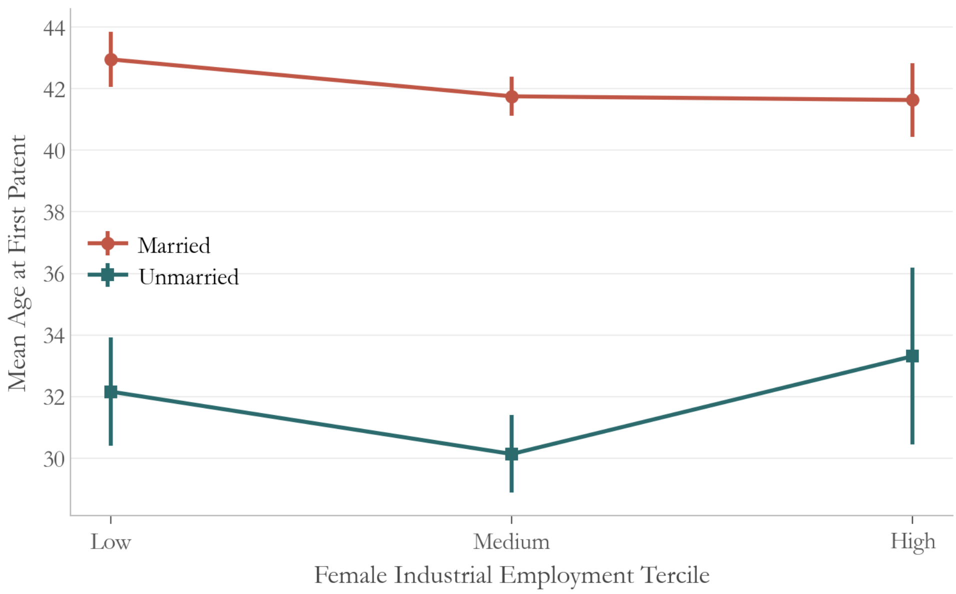


Figure C.2: Predicted age at first patent by marital status and female industrial employment tercile. The parallel lines confirm that the marriage gap is constant across levels of department-level industrialization.

C.9 Permutation distribution figure

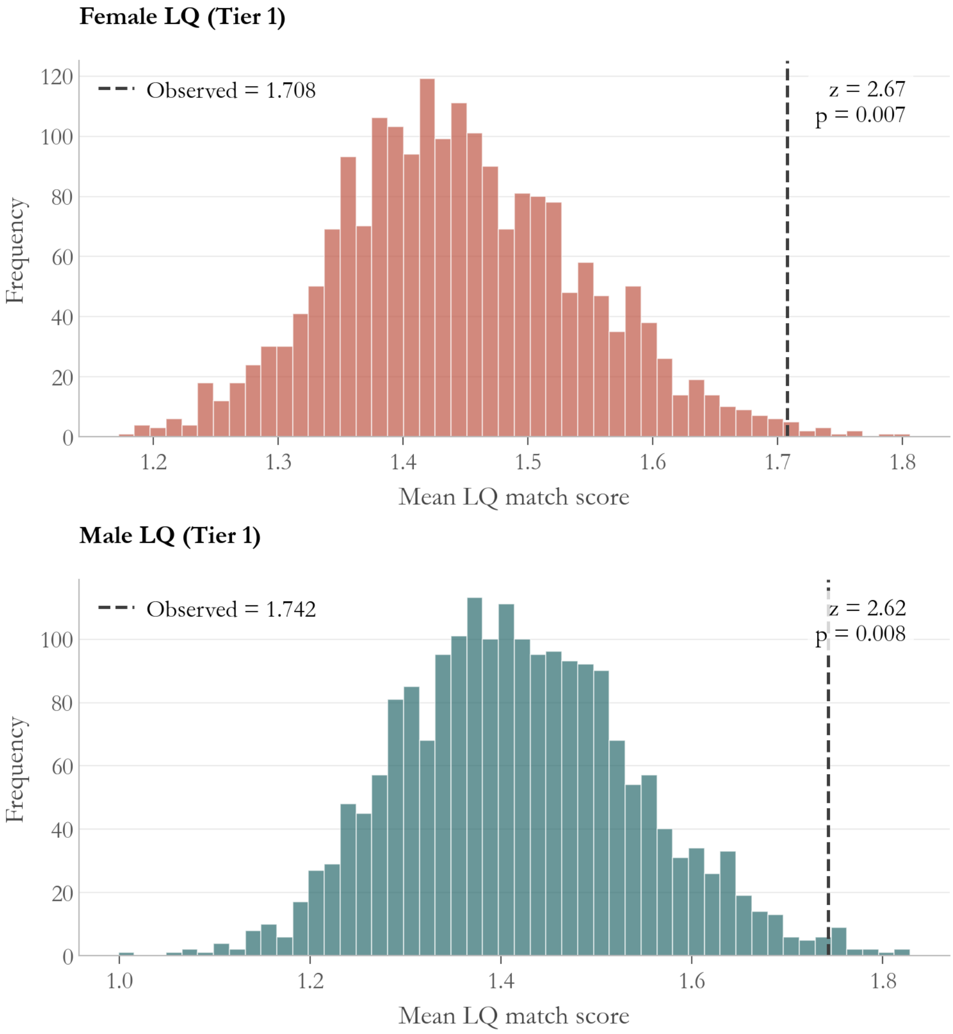


Figure C.3: Permutation distributions of mean Location Quotient match scores. Left: female employment. Right: male employment. Red dashed lines show observed values.

D Life cycle and childbearing

D.1 Supplementary life-cycle figures

These figures document the distributional detail behind the lifecycle findings in the main text, including the marriage-to-patent lag, the productivity split by starting age, and sectoral variation in patenting age.

Figure D.1: The innovation lag (married inventors only). Distribution of years elapsed between marriage and first patent (N = 659). Positive values indicate patenting after marriage. The median lag is 16.0 years, with 95.8 percent of inventors patenting post-nuptially.

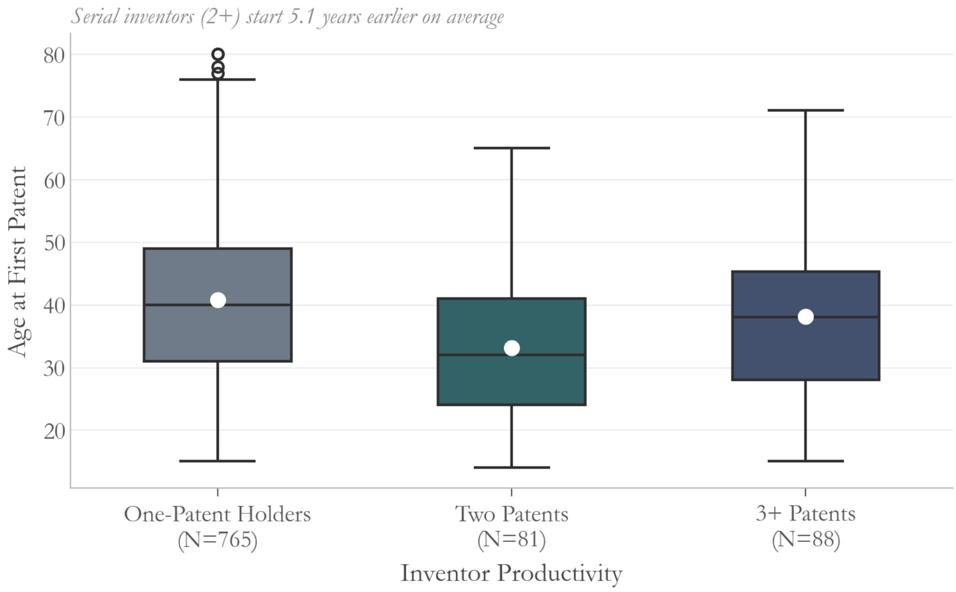


Figure D.2: Age at first patent by productivity class. Boxplots comparing starting age of one-time patent holders vs. serial inventors (2+ patents). Serial inventors filed their first patent substantially younger. Sample sizes shown on x-axis.

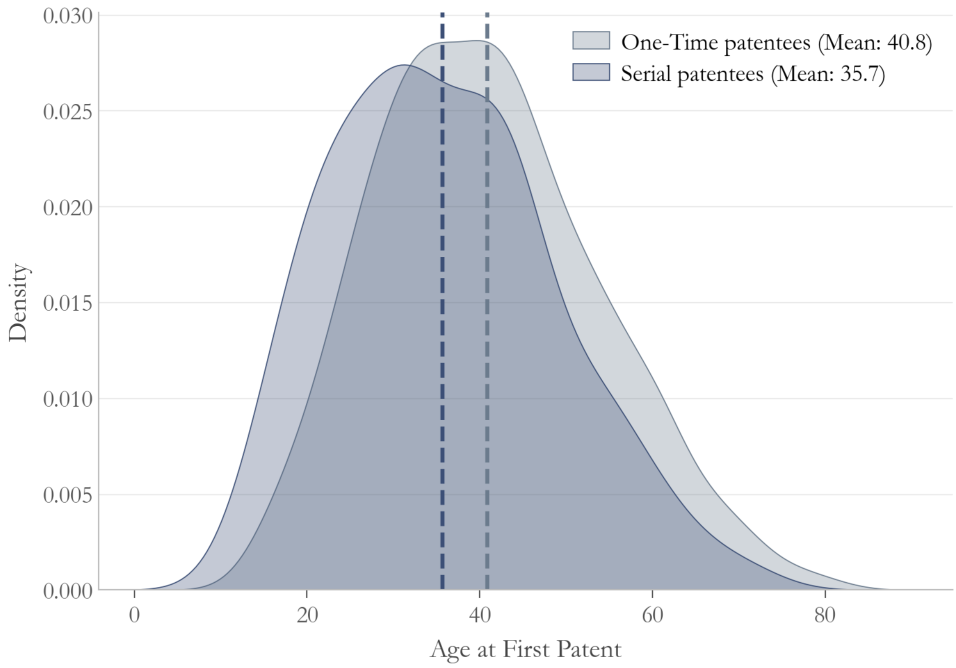


Figure D.3: Starting age density: one-time vs. serial patentees. Kernel density estimates show that the density peaks differ: serial inventors concentrate in the late twenties and early thirties, one-time patentees in the forties.

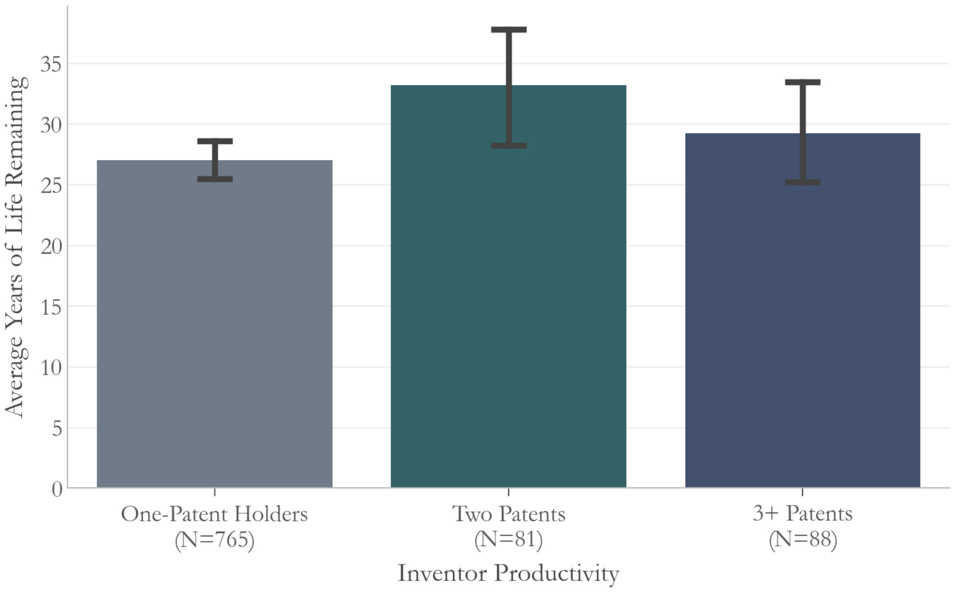


Figure D.4: Years lived after first patent by productivity class. Average remaining lifespan after first patent with 95 percent confidence intervals. Serial inventors lived longer after their first invention than one-time holders, consistent with their younger age at first patent.

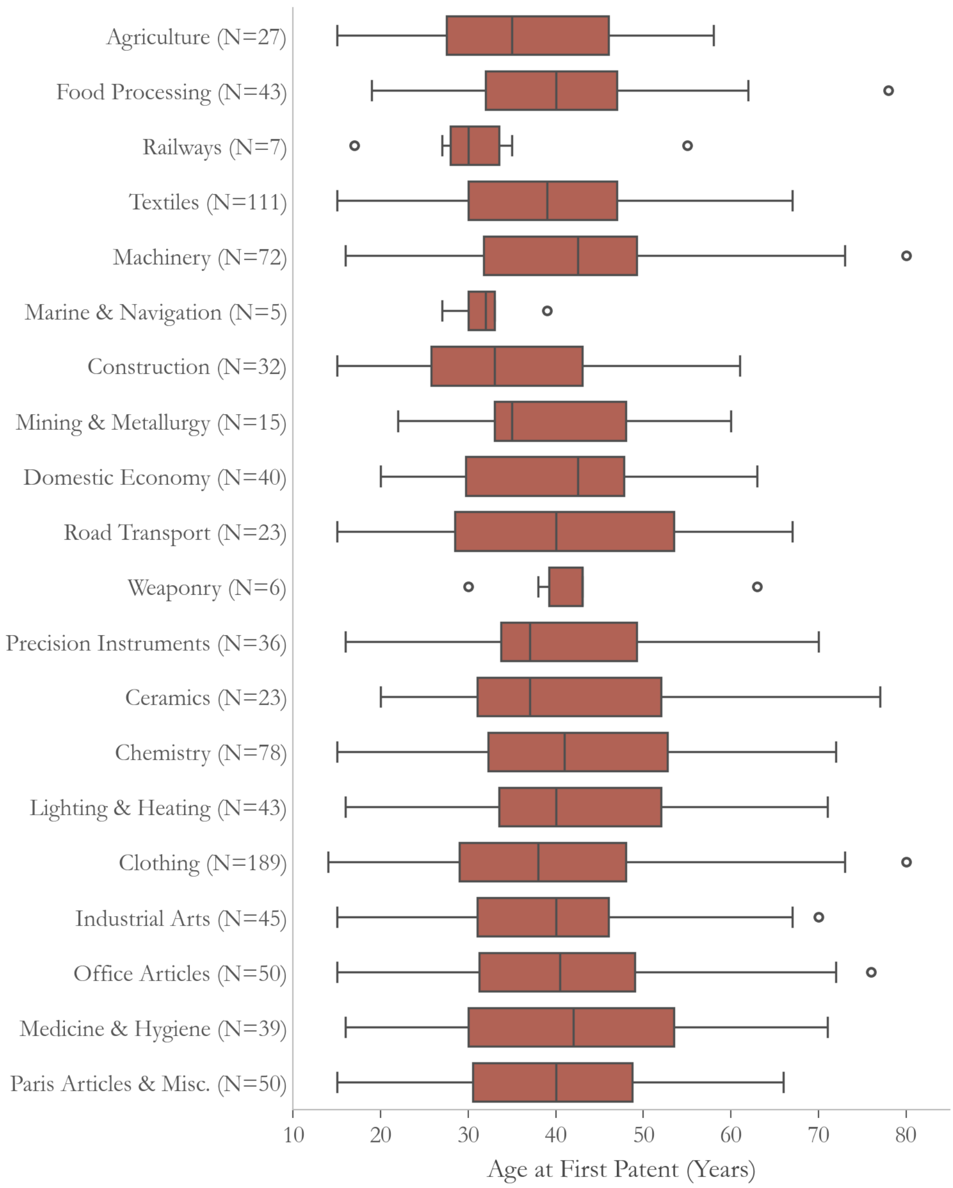


Figure D.5: Age at first patent by technology class. Boxplots of age distribution across 20 technology classes. Among sectors with at least 25 observations, Clothing shows the youngest profile and Chemistry the oldest, consistent with differences in human capital requirements and barriers to entry. Sample sizes shown on y-axis.

D.2 Supplementary childbearing figures

These figures document the fertility and timing patterns analyzed in the main text, including family size distributions, the childbearing window, the intersection of motherhood and invention, and the role of child mortality.

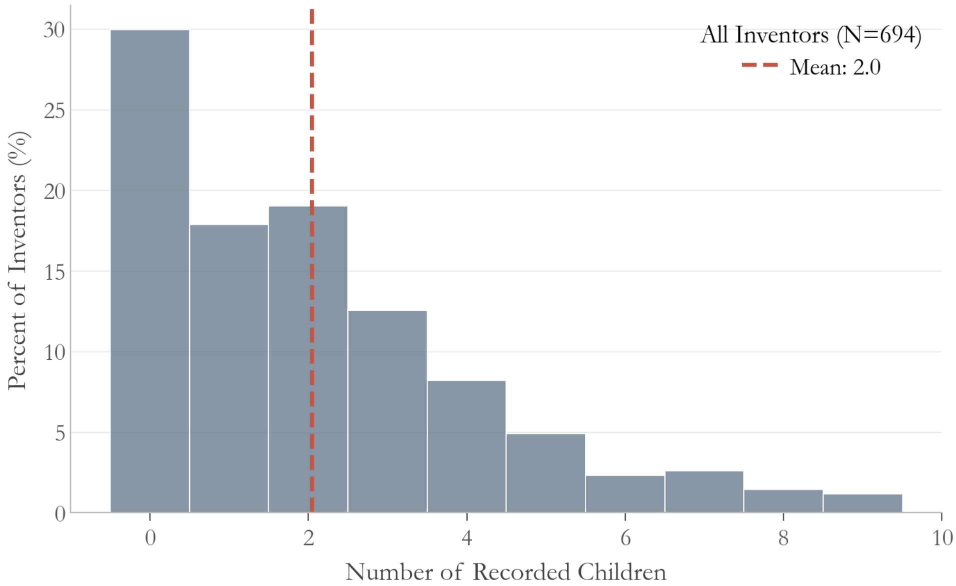


Figure D.6: Distribution of family size. Histogram of recorded number of children for all linked inventors (N = 694). The sample includes women with no recorded children (30.0 percent) and excludes outliers with more than 9 children to minimize measurement error.

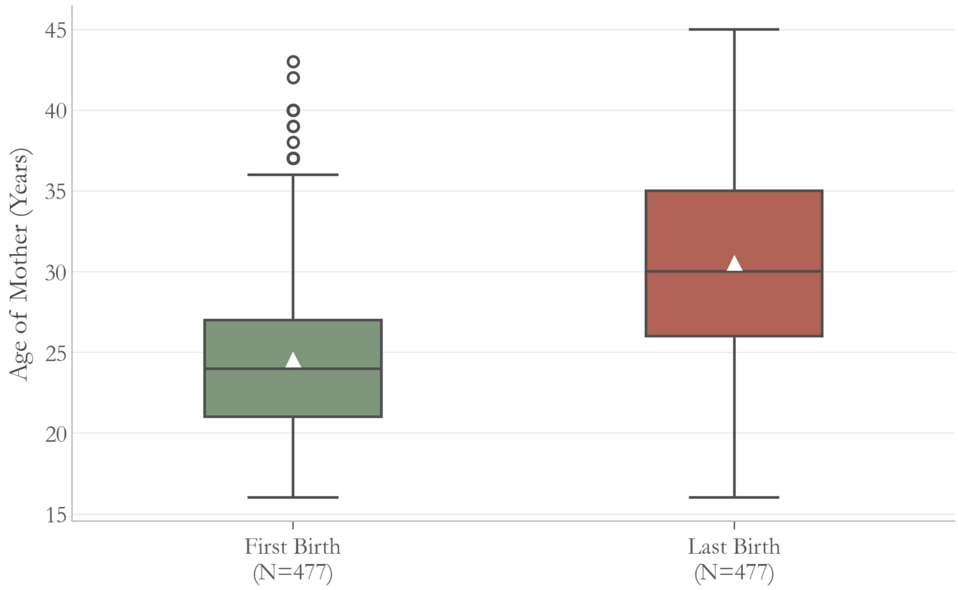


Figure D.7: The childbearing window. Distribution of maternal age at first and last birth for the sub-sample of mothers (N = 477). The average active childbearing span is 6.0 years.

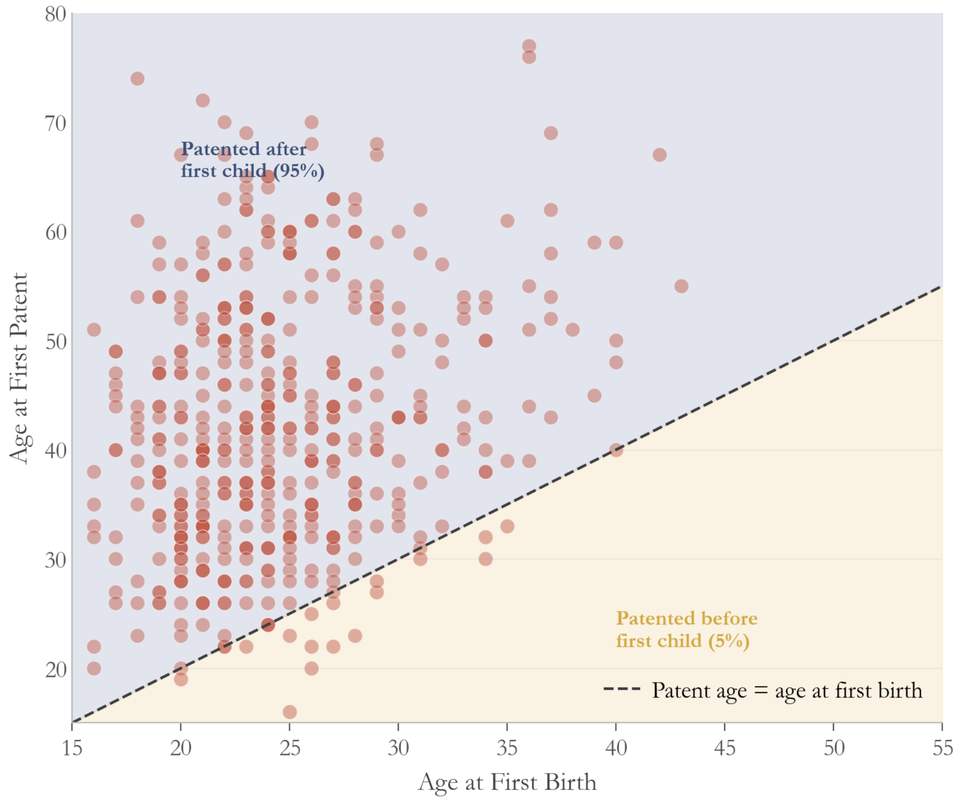


Figure D.8: Intersection of motherhood and invention. Scatter plot of age at first birth versus age at first patent. Points above the diagonal dashed line represent women who patented after becoming mothers (94.5 percent); points below indicate pre-maternal invention.

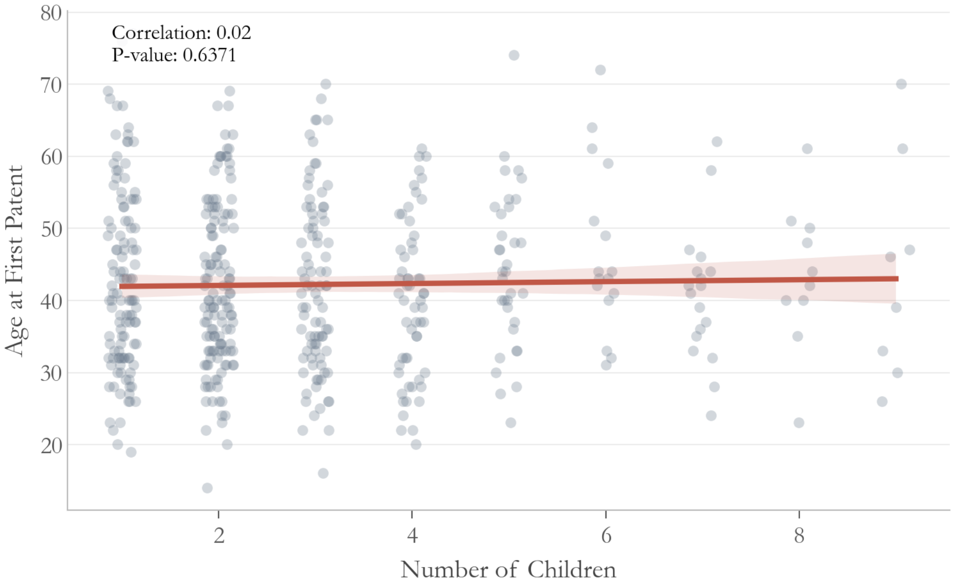


Figure D.9: Recorded family size and age at first patent. Scatter plot of recorded number of children versus age at first patent. The observed regression line ($r = 0.02$, $p = 0.64$) is close to flat, indicating that recorded family size was only weakly associated with when women patented in this sample.

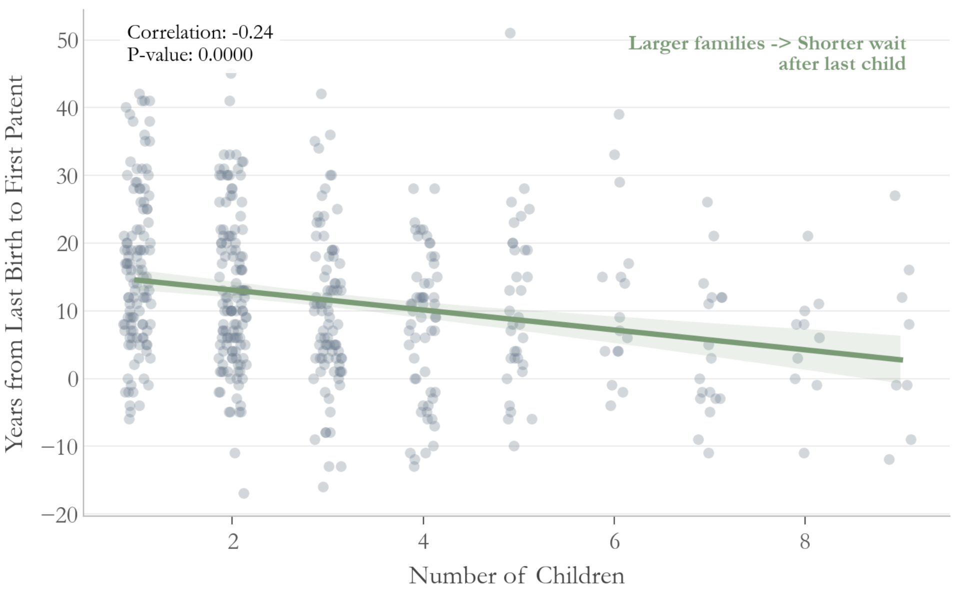


Figure D.10: Recorded family size and the post-childbearing gap. Relationship between recorded number of children and years elapsed from the last recorded birth to the first patent ($r = -0.24$, $p < 0.001$). The negative relationship is an arithmetic consequence of the observed pattern in Figure D.9: more recorded children push the last recorded birth later while the observed age at first patent remains roughly constant.

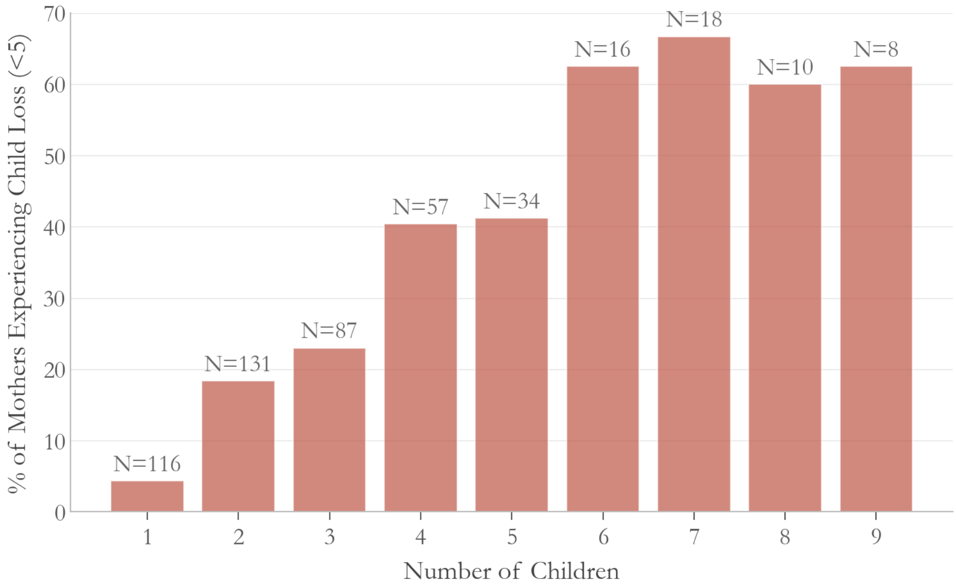


Figure D.11: Child mortality incidence by family size. The share of mothers who experienced at least one child death before age five, by number of recorded children. The incidence rises steeply with family size, from 4 percent among mothers of one child to over 67 percent among mothers of seven or more. N counts shown above each bar.

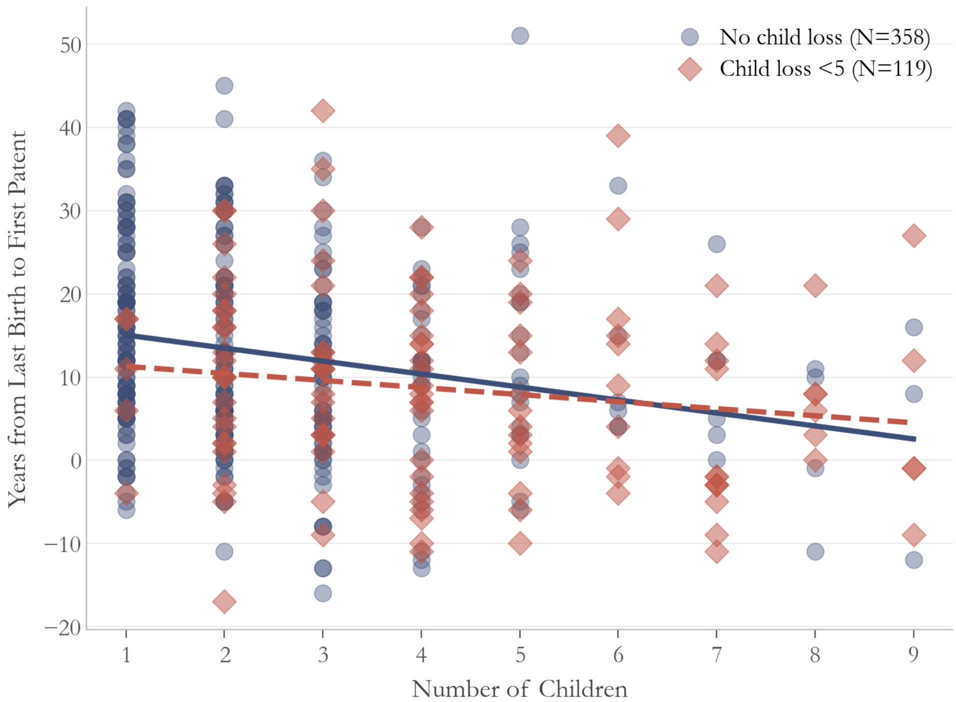


Figure D.12: Child mortality and the post-childbearing gap. The relationship between recorded family size and the gap from the last recorded birth to the first patent, with mothers separated by whether they experienced recorded child loss before age five. Both groups exhibit the negative slope that follows mechanically from the observed pattern in Figure D.9, though the no-loss group (blue, solid line) has a steeper decline. The gap is shorter for mothers who experienced recorded child loss at every family size, but this difference is not statistically significant once recorded family size is controlled.

D.3 Sectoral selection by family status

Recorded family size is only weakly related to patent counts (mean 1.42 for women with no recorded children, 1.39 for small families, 1.43 for large families; $p = 0.937$). The sample lacks power to rule out moderate effects, and recorded family size itself can understate completed fertility. Family status does, however, correlate with sector choice. Figure D.13 compares the sectoral distribution of first patents for mothers and women with no recorded children across the ten largest technology classes. Women with no recorded children concentrate in Clothing and Medicine & Hygiene (31 percent of their patents versus 20 percent for mothers), while mothers are overrepresented in Textiles, Chemistry, and Machinery (30 percent versus 24 percent).

These differences do not overturn the main finding that most women, regardless of family status, patented in mid-life. They do suggest that motherhood shifted the sectoral mix toward fields requiring more specialized technical knowledge. Whether this reflects self-selection, changes in household circumstances, or differences in access to production facilities remains open.

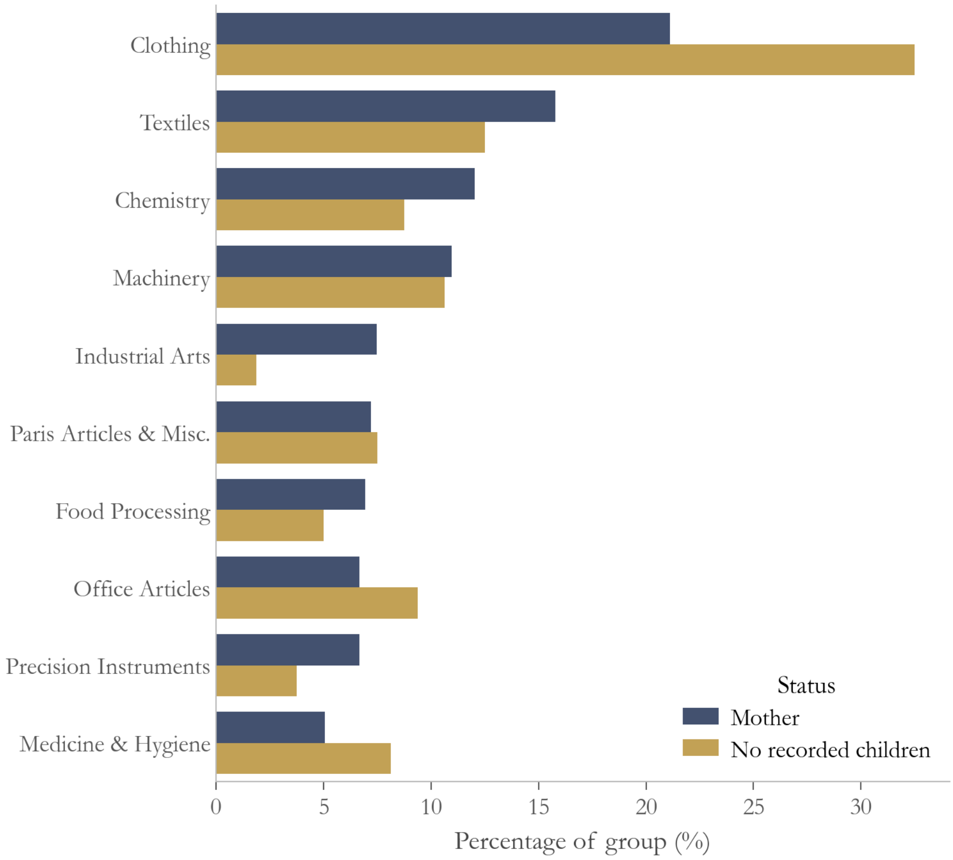


Figure D.13: Sectoral selection by maternal status. Comparison of the technological distribution of first patents for mothers versus women with no recorded children. Women with no recorded children are relatively more concentrated in Clothing and Medicine, while mothers are overrepresented in Textiles, Chemistry, and Machinery.

Table D.1: Pre/post-1835 cohort means in the age-capped balanced sample of mothers (N = 274)

Variable	Pre-1835 (n = 117)	Post-1835 (n = 157)	Difference
Age at first marriage	23.1	22.5	0.6
Age at last birth	31.2	29.1	2.1
Age at first patent	39.8	36.1	3.7
Gap (last birth to first patent)	8.7	7.0	1.7

Notes: Complete-case sample of mothers observable over ages 20 through 50 who filed their first patent by age 50. Differences are pre minus post, so positive values mean the pre-1835 mean is older.

D.4 Observation-window truncation and the corrected cohort comparison

The patent data cover inventions filed between 1791 and 1900. Women in the linked sample are born between 1764 and 1881. Because the analysis window ends in 1900, later-born cohorts are observed over a shorter portion of adult life than earlier-born cohorts. A woman born in 1820 can be observed filing a first patent at any age from 20 to 80; a woman born in 1870 can only be observed up to age 30. Raw cohort means therefore overstate the shift in age at first patent.

To correct for this, I restrict the comparison to mothers whose observation window covers ages 20 through 50 and who filed their first patent by age 50. Pre-1835 and post-1835 cohorts then contribute evidence over the same age range. Table D.1 reports the pre/post means of the life-cycle events relevant to patenting in this balanced sample.

The corrected comparison supports the narrower claim in the main text: within a common observation window, post-1835 mothers filed their first patents about four years earlier.