Becoming Cinderella: Strategic Beauty Investment Under Marriage Norms

Heng Chen* Xiangyu Lu* Qingxu Yang†

July 9, 2026

Abstract

This paper studies whether marriage-market incentives induce strategic investment in physical appearance. We exploit China's "leftover women" norm—a fertility-driven, discrete age threshold that penalizes women's marital desirability in their early thirties—as exogenous variation in beauty-investment returns. Using a nationwide inpatient dataset of over 100,000 plastic surgery admissions, we find that women's cosmetic procedures bunch sharply before this threshold, with no analogous pattern for functional procedures or men. Bunching concentrates in facial and breast procedures, intensifies with stronger fertility norms and marriage-market competition, and eases with women's labor-market opportunities. A search model rationalizes these patterns as a deadline effect.

Keywords: Beauty Premium, Premarital Investments, Fertility Norm, Bunching

*Faculty of Business and Economics, The University of Hong Kong. Email: hengchen@hku.hk.
*Faculty of Business and Economics, The University of Hong Kong. Email: lujeff@connect.hku.hk.
†School of Economics, Fudan University. Email: qxyang@fudan.edu.cn.

1 Introduction

Physical appearance generates substantial returns in the marriage market. Men weigh partner appearance far more heavily than women do (Fisman et al., 2006; Hitsch et al., 2010), granting attractive women preferential access to economically advantaged partners (Chiappori et al., 2017; Hamermesh, 2011). While the existing literature has largely treated beauty as an exogenous endowment, the technology for appearance enhancement has long been available and access to cosmetic procedures has widened substantially in recent decades, reaching an enormous economic scale. Yet remarkably little is known about whether individuals strategically invest in their physical appearance and what factors underlie their decisions.

Studying this question requires confronting two fundamental empirical challenges. The first is *measurement*: unlike most forms of investment, investments in beauty are inherently private, making such investments difficult to observe at scale. The second is *identification*: even when such investment is observed, the decision to undergo a procedure is shaped by a range of confounding factors. In the absence of exogenous variation in the returns to appearance, any observed correlation between cosmetic investment and economic outcomes remains susceptible to omitted variable bias.

We address both challenges by studying cosmetic surgery decisions in China, drawing on two complementary empirical components. The first is a large administrative dataset drawn from a healthcare system covering over 200 million inpatient records from more than 2,000 hospitals across China, from which we identify approximately 100,000 plastic surgery admissions between 2017 and 2019. Each record contains rich clinical information — including diagnosis codes, physician notes, operation records, and procedure-level classifications by body part — alongside a detailed demographic profile covering the patient's gender, age, region, and occupation, allowing us to study appearance investment at the individual-visit level.

The second component captures a salient social norm in the Chinese marriage market: the "leftover women" (*sheng nü*) norm. Chinese society—like other East Asian societies with similarly conservative marriage norms—holds a collective, informally enforced timetable for when a woman "ought" to marry. Women who remain unmarried past a socially prescribed age—typically 30—are labeled as such, perceived as lacking desirability in the marriage market and as a failure for both themselves and their family (Ji, 2015a; Liu, 2021; Fincher, 2023). Underlying this norm is a fertility-based logic: childbearing in China is expected to occur within marriage and is treated as a traditional obligation tied to family continuity, so a woman's marital desirability is evaluated largely through her perceived reproductive horizon, with the decline in fertility with age directly priced into the marriage market (Low, 2024b). As a woman grows older, her narrowing window for childbearing correspondingly diminishes her perceived desirability as a spouse (Gaetano, 2014; You et al., 2021).

As a consequence, this norm transforms a gradual biological decline in fertility into a discrete penalty in the marriage market: once an unmarried woman crosses the norm-induced mental threshold, she is stigmatized as a "leftover woman," and her marriage prospects deteriorate sharply. This penalty falls exclusively on women, with no comparable age threshold imposed on men. This asymmetric discontinuity offers plausibly exogenous variation in the returns to beauty investment, allowing us to identify how marriage-market incentives shape women's timing of cosmetic surgery.

A key input to our strategy is to locate the age threshold. Although the marriage norm literature and public discourse broadly place the penalty in a woman's early thirties, the precise threshold remains unspecified. We draw on data from the 2021 Chinese General Social Survey (CGSS), in which respondents rated the attractiveness of randomly constructed partner profiles varying in age and gender. Male respondents' ratings of female profiles decline discontinuously at age 31, while female respondents' ratings of male profiles exhibit no comparable pattern.[1]

With the threshold located, we turn to our main empirical question: whether this norm induces strategic beauty investment — specifically, whether women time cosmetic surgery to fall before the age-31 cutoff to marry before the penalty takes effect. If so, the age distribution of cosmetic procedures should exhibit excess mass just below the threshold relative to the smooth counterfactual.

To quantify this, we employ the bunching estimator of Chetty et al. (2011), which fits a smooth counterfactual to the observed surgery-age distribution and measures strategic retiming as the deviation from that baseline.[2] The resulting excess mass — the difference between observed and counterfactual surgery counts within that window, normalized by the counterfactual mean — directly quantifies the distortion in surgery timing attributable to the norm.

Our main findings establish that marriage-market incentives drive beauty investment decisions. First, we document substantial bunching in beauty-enhancing surgeries immediately before age 31: the pre-cutoff window contains 8.22 percent more surgeries than

[1] We further corroborate this cutoff using the method of Porter and Yu (2015), which endogenously detects a discrete kink in surgery records at age 31, with other individual characteristics remaining smooth across the threshold. The boundary also coincides with a well-documented biological constraint, as female fertility declines markedly after age 31 (Dunson et al., 2002, 2004).

[2] The method has been widely applied to settings where agents face sharp policy or social cutoffs, including labor supply responses to income tax kinks, teachers manipulating grades around passing thresholds (Dee et al., 2019), and firms timing R&D investment to qualify for policy subsidies (Chen et al., 2021).

the smooth counterfactual predicts, an economically substantial and precisely estimated response. Second, this bunching is specific to aesthetic procedures — functional surgeries exhibit no bunching around the threshold, mitigating concerns about age-specific supply-side or wealth-driven explanations. Third, we find no comparable bunching among men at any age, consistent with a marriage market in which the "leftover women" penalty applies exclusively to women.

Although our survey evidence identifies age 31 as the relevant threshold, social norms need not operate at a single precise age. Women may perceive the "leftover" stigma as applying broadly across their late twenties and early thirties, so strategic surgery timing could lie anywhere within this range. To verify that the observed bunching reflects a norm-induced response, we re-estimate the bunching coefficient at a range of placebo cutoff ages. Consistent with the diffuse nature of the norm, modest bunching emerges at ages immediately adjacent to 31 (notably 30 and 32), but excess mass peaks sharply at age 31, attenuates at neighboring ages, and vanishes at more distant placebo cutoffs. As a further robustness check, we pool ages 30–32 into a wider cutoff window, yielding an even larger bunching estimate.

Having established the robustness of our main result, we turn to its heterogeneity, examining when and for whom strategic beauty investment is more pronounced. First, if the response is driven by marriage-market incentives, it should concentrate in procedures most relevant to physical attractiveness in partner evaluation. Decomposing cosmetic admissions by body part, we find substantially larger excess mass for facial and breast surgeries—the features weighted most heavily in male partner evaluation—than for other aesthetic procedures. Women selectively retime the investments with the highest expected marital returns rather than advancing all cosmetic procedures uniformly.

Second, our interpretation hinges on a fertility-driven age deadline, so the response should be sharper where fertility norms are stronger. Classifying provinces by the local importance attached to bearing children to continue the family line, we find that bunching is far more pronounced where these norms are stronger. Another explanation is that fertility norms merely proxy for broader patriarchal values, so that the pattern reflects general conformity to gender norms rather than a specific age deadline. We test this alternative by exploiting the intensity of Confucian culture—a measure of patriarchal values that carries no built-in age threshold—and find no corresponding heterogeneity. Together, these results point to an age-specific marriage-market penalty rather than conservative gender norms per se.

Third, we turn to the role of marriage-market competition. Where competition for partners is fiercer, cosmetic procedures that improve women's matching prospects become more valuable, giving women stronger incentives to invest in their appearance before the cutoff. We capture this pressure with two complementary proxies: the ratio of active female to male users on Baihe.com, one of China's largest matchmaking platforms, where a higher ratio means women compete for a smaller pool of partners; and the frequency with which local residents search marriage-pressure terms such as "urged marriage" on Baidu, China's main search engine, which captures how strongly the pressure to marry is felt locally. Under both measures, bunching is stronger where marriage-market pressure is more intense.

Fourth, women's labor-market opportunities may substitute for appearance in marriage-market competition. Marital prospects can be enhanced either through beauty investment or through non-beauty attributes such as education, stable employment, and earnings; when labor-market returns are high, women rely more on these alternative channels and less on cosmetic surgery. Consistent with this substitution, bunching is markedly weaker among women with lower economic dependence, in regions with high female educational attainment, and where gender wage gaps are narrower—that is, precisely where women's economic alternatives are least constrained.

Finally, we examine a set of alternative interpretations—that the bunching reflects medical considerations, budget constraints, career-related incentives, or demand for postpartum recovery—and find that none is consistent with the observed patterns.

Together, these findings are organized and rationalized by a continuous-time marriage-market search model. A single woman searches for a partner over her life cycle and chooses when to undertake a one-time costly investment that permanently raises her matching rate. Her timing balances two forces: the marginal benefit of investing earlier — the incremental matching hazard from enhancement multiplied by the marriage surplus — and the marginal cost, comprising the opportunity cost of the expenditure and the expected cost of a procedure that proves unnecessary if she matches naturally before investing.

Absent the age-based stigma, each woman balances these two forces to choose her optimal surgery age, and the resulting distribution of surgery timing across the population is smooth. When the norm is present, it imposes a discrete penalty at an age threshold, reducing the continuation value of remaining single beyond it and widening the pre-threshold marriage surplus. This tilts the trade-off toward earlier investment and generates a deadline effect: women whose unconstrained optimal timing lies just above the cutoff advance their procedure to strengthen marriage prospects before the age penalty takes effect. Across the population, this forward substitution generates excess mass in the surgery-age distribution immediately below the threshold — the bunching pattern we document — and a corresponding deficit just above it.

The model also delivers comparative statics that align with the heterogeneity patterns in the data. First, bunching intensifies when the marriage-market return to advancing beauty investment is higher — either because appearance plays a larger role in partner evaluation, because the penalty attached to crossing the fertility deadline is sharper, or because competition for partners makes improvements in matching prospects more valuable. In the data, this is reflected in the larger bunching for face and breast procedures — the features most scrutinized in partner evaluation — in regions where fertility norms are stronger, and in regions where marriage-market competition is fiercer. Second, bunching weakens when women have stronger labor-market opportunities: because marital appeal can be built through both beauty and non-beauty investments, higher returns to education and employment induce substitution away from cosmetic surgery. The data confirm this: bunching is weaker among women with lower economic dependence, in provinces with higher female educational attainment, and where gender wage gaps are narrower. Conversely, women lean more heavily on beauty investment when their labor-market opportunities and the returns to non-beauty investment are limited.

We now relate our findings to several strands of literature in economics. First, we contribute to the literature on fertility by showing that fertility-based gender norms shape women's choices well before childbirth. A large body of work documents a substantial and persistent penalty for women that emerges *after* the birth of their first child and operates primarily through reduced labor-market performance (Goldin, 2014; Kleven et al., 2019, 2025). More recent work shows that part of this penalty is anticipatory and rooted in fertility itself: women of reproductive age are penalized for their *potential* motherhood even in the absence of any childbirth, with the penalty fading once they age out of the fertile window (Petit, 2007; Adda et al., 2017; Bursztyn et al., 2017; Zamberlan and Barbieri, 2023).

We extend this literature by pushing its central margin earlier in the life cycle. By tying women's marital desirability to their perceived reproductive horizon, the "leftover women" norm turns a gradual decline in fertility into a socially salient age deadline—one that operates in the *marriage* market, before marriage and childbearing, rather than in the workplace afterward. This deadline induces women to undertake costly and irreversible cosmetic-surgery investments, a pre-birth behavioral response that, to our knowledge, has not previously been documented.

Second, we contribute to the economics of beauty. A large body of work documents that physical attractiveness generates returns across economic domains: in labor markets through higher wages and better hiring outcomes (Hamermesh and Biddle, 1994; Mobius and Rosenblat, 2006; Kuhn and Shen, 2013); in corporate settings through higher executive compensation (Graham et al., 2017); and, most starkly, in the marriage market, where attractive women gain preferential access to economically advantaged partners (Fisman et al., 2006; Hitsch et al., 2010; Chiappori et al., 2012; Zhang et al., 2023). Yet this literature has largely treated attractiveness as a fixed, exogenous endowment that shapes outcomes but is not itself a matter of choice (Hamermesh, 2011). We recast it as an endogenous investment: women actively produce beauty when doing so pays off. In our setting, what makes it pay is the marriage market's premium on a youthful, fertile appearance—so that the same fertility-based norm that penalizes women through the workplace also drives them, through the marriage market, to invest in their looks.

Finally, we also add beauty as a previously overlooked channel to the literature on premarital investment (Peters and Siow, 2002; Chiappori et al., 2009; Bhaskar and Hopkins, 2016; Low, 2024a), including in China, where marriage-market forces are known to shape savings, intergenerational transfers, and matching outcomes (Wei and Zhang, 2011; Han and Shi, 2019; Bhaskar et al., 2023; Han et al., 2022, 2026).[3]

2 Background

2.1 The Cosmetic Surgery Market

Cosmetic surgery has been practiced in various forms since the early twentieth century, with modern techniques developing notably during and after the two World Wars as reconstructive surgery advanced. Over subsequent decades, procedures expanded from reconstructive to elective cosmetic applications, and by the second half of the twentieth century, operations such as rhinoplasty and breast augmentation had become widely performed. Advances in anesthesia, minimally invasive techniques, and outpatient care have since made cosmetic surgery progressively safer, less costly, and accessible to a broad population. This expanding accessibility underlies the rapid growth observed in recent years.

The global cosmetic surgery market has grown substantially over the past decade and now constitutes a major segment of the broader beauty economy. Worldwide, 34.9 million aesthetic procedures were performed in 2023, a 42 percent increase over four years (ISAPS, 2023). The global market was valued at approximately USD 83 billion in 2024 and is projected to double by 2033 (Grand View Research, 2024), reflecting both rising demand and expanding access across income levels.

Within China, the consumer base is heavily concentrated among young women. Approximately 90 percent of cosmetic surgery consumers are women under the age of 35, and China has emerged as one of the largest national markets and a central driver of global expansion (Wu et al., 2022). The most frequently performed procedures target facial features, with double-eyelid surgery, rhinoplasty, and facial contouring among the most common. Providers

[3]Our paper also relates to a broader literature on how marriage and gender norms distort women's economic decisions absent formal constraints (Bertrand et al., 2015; Bursztyn et al., 2020; Ashraf et al., 2020; Jayachandran, 2021).

range from hospital-affiliated aesthetic departments to a rapidly expanding network of private specialty clinics concentrated in major urban centers. This concentration of cosmetic surgery demand among young women approaching marriage age is precisely the pattern this paper seeks to explain.

2.2 "Leftover Women" Norm

The composition of this demand — concentrated among unmarried women in their late twenties and early thirties — points to the marriage-market pressures we examine next.

China's "leftover women" (*sheng nü*) norm stigmatizes women who remain unmarried beyond a mental threshold—a socially prescribed age that frames delayed marriage as a familial failure rather than a personal choice (To, 2013; Ji, 2015b; You et al., 2021). Age is already a salient attribute in marriage-market evaluations, and the norm further amplifies its importance: women who remain unmarried into their late twenties and early thirties face a shrinking pool of potential matches and heightened pressure to marry before crossing the age threshold.

What makes this norm particularly consequential is its specifically age-based character, driven by two reinforcing forces. First, female fertility declines with age — the so-called "biological clock" (Navot et al., 1991; Broekmans et al., 2007) — lending urgency to the timing of marriage and childbearing. Second, this fertility decline carries particular weight in China, where marriage is closely embedded in family continuity and intergenerational obligation. Parents and relatives frequently participate in partner search, and because marriage and childbearing are tied to family succession, a daughter's delayed marriage becomes a source of stigma for the entire family.[4] Delayed marriage thus carries a cost not just for the individual but for the family's intergenerational continuity.

At its core, the fertility-based norm concerns reproductive timing rather than age as such: because marriage and childbearing are inseparable in the Chinese family context, a woman's marriage-market value is tied to her perceived capacity to bear children. Indeed, the very emergence of the *sheng nü* label in public discourse was bound up with anxieties over female fertility and reproduction (Liu, 2021; Fincher, 2023).

Despite the clear social force of this norm, its precise age threshold is not formally defined. The "leftover women" discourse is widely recognized, but the exact age at which the penalty applies remains unspecified. Unlike a legal rule, the norm has no formally announced cutoff; public discussion frequently points to the late twenties or early thirties,

[4]Similar links between women's marriage timing, family honor, and social stigma have been documented in other developing-country contexts, particularly in South Asia; see, for example, Nanda et al. (2012) and Malhotra and Elnakib (2021).

but no sharp age is explicitly defined. One plausible candidate is the early thirties: medical evidence shows that female fertility declines more steeply from the early thirties onward (Broekmans et al., 2007), and in a social environment that places strong value on marriage, childbearing, and intergenerational continuity, this biological gradient may render the early thirties a particularly salient threshold in China's marriage market. Identifying this threshold empirically is therefore central to the analysis that follows.

3 Data

3.1 Data Source

Our primary data source is a large Chinese healthcare information system that records inpatient encounters from partner hospitals nationwide. The raw database contains over 200 million inpatient records from more than 2,000 hospitals. For each hospitalization, we observe patient identifiers, age, gender, occupation, admission date, length of stay, total expenditure, ICD-10 diagnosis codes, diagnosis notes, and operation records. Together, these variables allow us to characterize both the timing and the clinical content of plastic-surgery admissions.

All observations are inpatient records and therefore involve at least one overnight stay.[5] Diagnosis codes capture physicians' formal coding of each case, while diagnosis notes provide more granular detail through standardized entries or physician-entered free text. Operation records document the procedures performed during the hospitalization. Combining these sources enables us to identify plastic-surgery cases and to distinguish cosmetic procedures from medically indicated reconstructive ones.

We supplement the hospital records with several additional datasets. For marriage norms, we use the 2010 China Family Panel Studies (CFPS)—respondents' reported importance of having children for continuing the family line, proxying fertility-norm intensity, and prefectural clan density, proxying the strength of Confucian culture, following Greif and Tabellini (2017). For marriage-market pressure, we use Baihe.com user data—the active female-to-male ratio, proxying local competition—and the Baidu Search Index for marriage-pressure keywords, proxying perceived urgency. For women's economic opportunities, we use the 2015 one-percent Population Census for female educational attainment and industry wage statistics for regional gender wage gaps.

[5] The data therefore capture the more intensive segment of the plastic-surgery market, such as facial bone contouring or extensive liposuction, rather than minimally invasive outpatient treatments such as Botox injections. This limits the scope of our analysis but is useful for studying deliberate investment decisions, since inpatient procedures are typically costly, risky, and require nontrivial recovery time.

3.2 Sample Construction

We construct the analysis sample from inpatient records spanning 2017 to 2019.[6] Starting from all hospitalizations in this window, we identify plastic-surgery admissions using ICD-10 diagnosis codes together with accompanying clinical descriptions. We retain records whose primary or secondary diagnosis explicitly involves body-part reshaping, cosmetic alteration, or implant placement.

We then impose four restrictions. First, when a patient has multiple records in the same year, hospital, and diagnosis category, we retain only the initial admission. Second, we restrict the sample to individuals aged 10 to 60. Third, we exclude childbirth-related diagnoses, including genital-tract repair during or after childbirth. Fourth, we retain only records that can be linked to a specific plastic-surgery episode through diagnosis codes, clinical notes, or operation records; ambiguous records with plastic-surgery-related coding but no identifiable focal procedure are dropped. The resulting sample contains 102,699 plastic-surgery admissions from 617 hospitals across 216 prefecture-level cities.

We further classify admissions by medical purpose. Our baseline distinction is between *cosmetic* procedures, undertaken primarily for aesthetic enhancement, and *functional* procedures, performed to correct conditions that impair physical functioning or quality of life. Classification draws on both diagnosis codes and notes. Some ICD-10 descriptions are directly informative—for example, Z41.1 denotes encounters for cosmetic surgery. More generally, an admission is classified as cosmetic when the recorded procedure is primarily intended to improve appearance, and as functional when the primary purpose is reconstructive or medically indicated.[7] Appendix B provides the detailed coding rules.

3.3 Summary Statistics

Appendix Table A2 summarizes the plastic-surgery sample. Admissions are nearly evenly split by gender, with 51,726 male and 50,973 female admissions, but their composition differs sharply. Among women, 62.9% of admissions are classified as *For Beauty*, compared with 31.8% among men. Conversely, men are more likely to undergo functional procedures: 60.5% of male admissions are *For Functionality*, versus 26.5% among women.

[6] We begin in 2017 because hospital entry into the database is sparse beforehand and stabilizes thereafter. We end before 2020 to avoid COVID-19 disruptions to hospital operations and elective procedures.

[7] Cosmetic examples include breast augmentation, aesthetic rhinoplasty, double-eyelid surgery, and facial contouring. Functional examples include cleft-lip repair, burn reconstruction, and deformity correction. A small residual group—such as ptosis or breast-hyperplasia cases whose treatment involves reshaping but whose primary indication is not aesthetic—is classified as "Other," excluded from the baseline estimations, and retained for robustness checks.

The age distribution also differs markedly by gender. Male admissions are concentrated among adolescents, who account for 49.1% of the male sample. Female admissions, by contrast, are concentrated in the adult ages most relevant to our analysis: 25.7% are in their twenties, 21.2% in their thirties, and 18.7% in their forties.

Eye-related surgeries are the most common procedure category for both genders, accounting for roughly one quarter of admissions. Beyond this, body-part composition differs substantially. Breast procedures account for 22.3% of female admissions but only 6.7% of male admissions, whereas genital procedures account for 18.4% of male admissions and almost none of the female sample. Figure A1 decomposes admissions by body part and purpose, showing that cosmetic procedures dominate among women while functional procedures dominate among men.

Finally, the inpatient procedures we observe are relatively intensive. The average length of stay is 7.68 days, and mean spending per admission is 11,871 RMB. Cosmetic procedures are more expensive than functional ones, averaging 14,306 RMB compared with 9,535 RMB. These patterns reinforce that our data capture costly procedures rather than minor outpatient aesthetic adjustments.

4 Estimating Strategic Beauty Investment

4.1 Identifying the "Leftover Women" Age Threshold

Although it is widely believed that women are labeled as "leftover" after reaching a certain age, the precise threshold remains unclear. To examine whether there exists a specific age at which women's perceived marital desirability declines discretely, we draw on data from the 2021 Chinese General Social Survey (CGSS). In this survey, each respondent evaluated three profiles of hypothetical marriage partners. Profile attributes were randomly assigned along six dimensions—age, appearance, housing, income, family background, and education—and respondents rated each profile's attractiveness on a 7-point scale. Because the attributes of the hypothetical partners are randomly assigned, this design isolates how attractiveness evaluations vary with age, holding all other profile characteristics orthogonal by construction (see Table A1 for details of the profile randomization).

We begin by examining how attractiveness ratings vary with the partner's assigned age, after partialling out the effects of the remaining profile attributes and respondent characteristics. Figure A2 plots the residual against the assigned age of the hypothetical partner, separately for male and female respondents. The left panel, based on male respondents evaluating female profiles, reveals a clear downward shift around age 31: ratings remain broadly stable at younger ages but decline discretely once the female profile crosses this threshold. The right panel, based on female respondents evaluating male profiles, shows no comparable discontinuity at the same age. These patterns indicate that age enters partner evaluation asymmetrically, with a substantial penalty for women.

To locate the threshold more formally, we conduct a search over a set of candidate cutoff ages and identify the point at which the decline in women's evaluated attractiveness is largest. We restrict attention to the range between ages 27 and 35, which spans the ages most commonly associated with delayed marriage and the "leftover women" discourse in China (Graham, 2015). For each candidate cutoff, we compare the average residualized attractiveness score in a local window on either side of that age, and select the value that maximizes the estimated discontinuity. This procedure follows the logic of cutoff estimation when the threshold is not known ex ante (Porter and Yu, 2015). The resulting estimate is age 31, confirming the pattern evident in Figure A2 and consistent with a salient biological-clock constraint, as female fertility declines markedly beyond this age (Dunson et al., 2002, 2004). With this threshold established, we turn to estimating whether it distorts the timing of cosmetic surgery.

4.2 Empirical Strategy: Bunching Estimator

Our key hypothesis is that women facing a discrete drop in marriage-market prospects at the "leftover women" threshold will optimally retime cosmetic surgery to just before the cutoff. Taking age 31 as the baseline threshold identified in Section 4.1, this retiming produces an observable empirical pattern: excess mass in the age distribution of cosmetic-surgery admissions in the years immediately preceding age 31, relative to a counterfactual in which the distribution would otherwise evolve smoothly. Our goal is to estimate the magnitude of this excess mass.

Interpreting excess mass as evidence of strategic retiming rests on a key identifying assumption: absent the age norm, the age distribution of cosmetic surgery would evolve smoothly through age 31. Equivalently, no other determinant of surgery timing—whether biological, economic, or institutional—generates a discontinuity at this age.

Two features of our setting support this assumption, both documented below. First, patient composition does not change discontinuously at the threshold: observable characteristics including education, employment status, income, and housing evolve smoothly through age 31 (Figure A3). Second, no comparable bunching appears on placebo margins — neither for functional surgeries, male patients, nor at alternative cutoff ages (Section 4.3) — isolating the effect to marriage-market incentives specific to women.

To implement this design, we estimate the age distribution of surgery admissions separately by gender and by procedure type, using the bunching approach introduced above. Let F_a denote the share of first plastic-surgery admissions occurring at age a within a given estimation sample. We estimate the counterfactual age distribution by fitting a flexible polynomial to the observed age profile while excluding a window around the cutoff:

$$F_a = \sum_{q=0}^{Q} \beta_q a^q + \sum_{j=\hat{a}-h}^{\hat{a}+h} \lambda_j \mathbf{1}[a = j] + \varepsilon_a, \tag{1}$$

where $\hat{a} = 31$ is the cutoff age, Q is the polynomial order, and the indicator variables $\{\lambda_j\}$ flexibly absorb deviations of the observed age distribution from the smooth polynomial trend within the excluded window. In the baseline specification, we use one-year age bins, a seventh-degree polynomial, and a half-window of $h = 3$ years. The estimated counterfactual is then given by the fitted polynomial evaluated without the cutoff-window indicators:

$$\hat{F}_a = \sum_{q=0}^{Q} \hat{\beta}_q a^q.$$

We summarize the degree of bunching using the normalized excess mass in the pre-cutoff window:

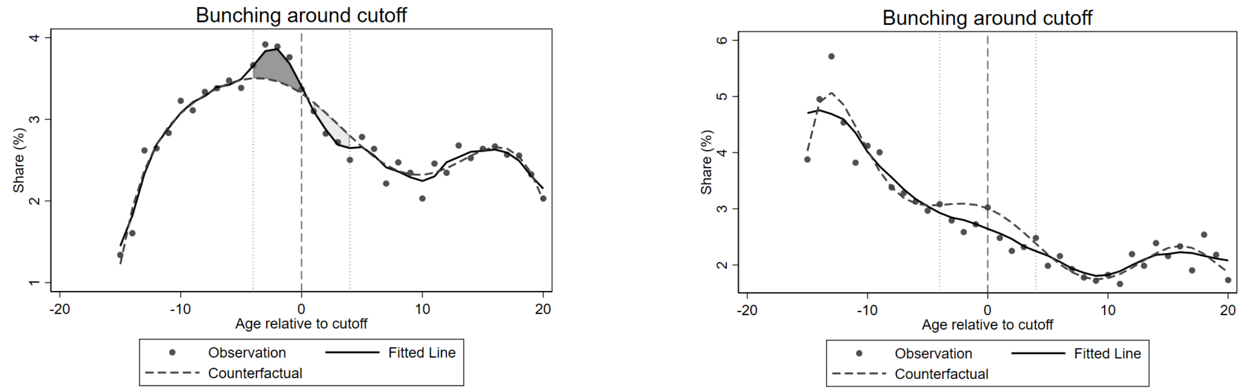
$$\hat{B} = \frac{\sum_{i=\hat{a}-h}^{\hat{a}} \left(F_i - \hat{F}_i \right)}{\sum_{i=\hat{a}-h}^{\hat{a}} \hat{F}_i} = \frac{\sum_{i=\hat{a}-h}^{\hat{a}} \hat{\lambda}_i}{\sum_{i=\hat{a}-h}^{\hat{a}} \hat{F}_i}. \tag{2}$$

This statistic measures the percentage excess of observed surgeries in the pre-cutoff window relative to the smooth counterfactual, and thus summarizes the extent of bunching around the cutoff in a given estimation sample. A positive value of $\hat{B}$ indicates that surgeries are disproportionately concentrated at ages just below 31 relative to the no-norm benchmark.

Standard errors are computed using the parametric bootstrap procedure of Chetty et al. (2011). Specifically, we resample from the estimated residual distribution of equation (1), reconstruct the implied age distribution, and re-estimate $\hat{B}$ in each replication. The reported standard error is the standard deviation of $\hat{B}$ across 1,000 bootstrap replications. We now apply this estimator to the data.

4.3 Results: Strategic Beauty Investment

Women's strategic beauty investment. We construct the age distribution of female plastic surgery admissions by aggregating individual-level records into one-year bins and



For Beauty, $\hat{B} = 8.22^{***}(1.34)$ **For Functionality**

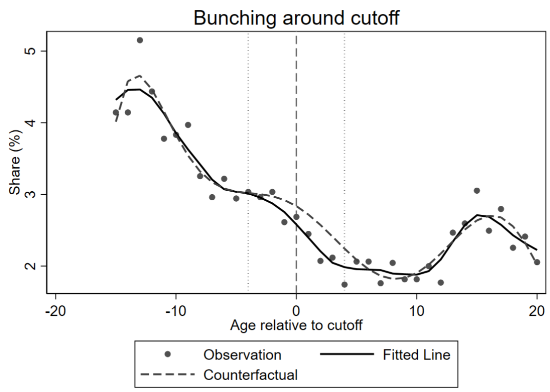
Figure 1: Female Age Distribution around the Age-31 Cutoff

Notes: This figure plots female patients' age-at-first-surgery distributions around the age-31 cutoff, separately for beauty-related and functionality-related procedures. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. In the left panel, the dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses. The right panel reports the corresponding distribution for functionality-related procedures.

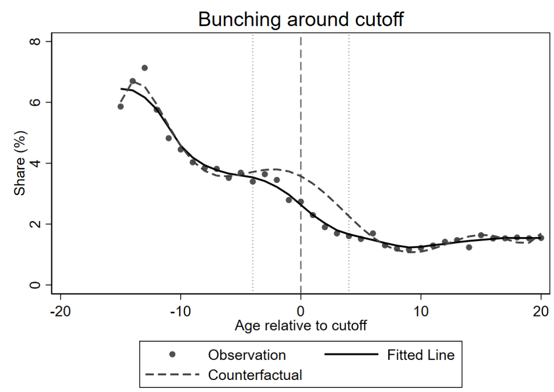
recentering each observation relative to the age-31 cutoff, following the procedure outlined in Section 4.2. Figure 1 plots the observed age distribution (dots), the fitted age profile (solid line), and the estimated counterfactual (dashed line) from Equation (1).

The left panel restricts the sample to beauty-enhancing surgeries. The empirical distribution exhibits a pronounced spike immediately to the left of the cutoff and a corresponding deficit to the right—the characteristic bunching signature of a discrete penalty at an age threshold. The dark shaded region marks the excess mass just below the cutoff, and the light shaded region the corresponding missing mass just above it. Consistent with our hypothesis, these two masses are linked by forward substitution: women whose unconstrained optimal surgery age falls just above age 31 rationally advance their procedure to strengthen their marriage prospects before the age penalty takes effect. We estimate a bunching coefficient of $\hat{B} = 8.22$ (SE = 1.34), indicating that cosmetic admissions in the pre-cutoff window exceed the smooth counterfactual by 8.22 percent—an economically substantial distortion attributable to the leftover-women norm.

Functional surgeries as a within-sample placebo. The right panel of Figure 1 plots the corresponding distribution for functional, non-cosmetic procedures. Unlike the left panel, it evolves smoothly through the age-31 threshold with no detectable excess mass in the pre-cutoff window. This contrast mitigates concerns about two natural alternatives:



For Beauty



For Functionality

Figure 2: Male Age Distribution around the Age-31 Cutoff

Notes: This figure plots male patients' age-at-first-surgery distributions around the age-31 cutoff, separately for beauty-related and functionality-related procedures. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. In contrast to the female cosmetic-surgery distribution, neither panel displays a clear excess mass immediately to the left of the cutoff, consistent with the absence of a comparable age-based marriage-market penalty for men.

age-specific supply-side shocks—such as differential physician referrals or hospital exposure around age 31—would generate a discontinuity in medically indicated admissions as well, and a demand-side health shock unrelated to marriage-market incentives would produce the same pattern in non-aesthetic procedures. Neither appears, and the excess mass in the cosmetic distribution therefore reflects deliberate, aesthetically motivated behavior rather than a general age-related shift in medical utilization.

Men's absence of response as a cross-gender placebo. Men provide an independent cross-gender placebo test. In China's marriage market, men's desirability is determined primarily by economic resources and social status rather than physical appearance, and men face no age-specific norm analogous to the leftover-women stigma. Figure 2 confirms this prediction: both the cosmetic and functional age distributions for men are smooth through the threshold, with no statistically significant excess mass in either panel.

The absence of a male response mitigates gender-neutral drivers of cosmetic demand: were the bunching due to a common, non-marital concern with appearance around the early thirties, it should appear among men as well, yet it does not. The response is therefore specific to women, as the marriage-market interpretation requires.

The gender asymmetry reveals an alternative channel through which social norms generate inequality. Because the age penalty falls exclusively on women, the full costs of strategic forward investment, comprising financial expenditure, surgical risk, and recovery time, are borne entirely by them. The bunching pattern documented here thus constitutes a form of implicit "pink tax" imposed by the leftover-women norm, a welfare cost falling on women in response to a socially imposed discontinuity in their marriage-market standing.

Robustness checks. A potential concern is that the estimated bunching coefficient is sensitive to the parameter choices in Equation (1), specifically the polynomial order Q and the half-window h. Figure A4 addresses this concern by presenting estimates across a range of polynomial orders and bandwidth values, and the estimated excess mass is stable throughout. Figure A5 shows that a pronounced bunching pattern persists when all female surgery cases, both cosmetic and functional, are pooled into a single sample. A remaining question, however, is whether the response is tied specifically to age 31 or reflects a broader behavioral shift across adjacent ages — a concern we address in the next subsection.

4.4 Norm-induced Strategic Investment.

Much of the bunching literature studies behavioral responses to explicit policy cutoffs — tax kinks, grade thresholds, or regulatory notches — where the discontinuity is publicly known and precisely defined (Dee et al., 2019). The "leftover women" norm is different in a crucial respect: it is a social norm, not a policy rule, and carries no officially announced age threshold. Social norms operate through beliefs about how others evaluate and respond to one's behavior, and such perceived beliefs can shape behavior even when they are imperfectly measured or partially misperceived (Bursztyn and Yang, 2022). Women do not face a sharp, known kink; rather, they perceive a diffuse stigma that intensifies across the late twenties and early thirties (Graham, 2015). The normative penalty may therefore operate over a range of ages rather than at a single point, raising the concern that the bunching we document at age 31 may reflect a broader response spanning adjacent ages.

We assess this by estimating the bunching magnitude using each candidate age from 27 onward as the cutoff. Specifically, we re-apply our bunching estimators around each candidate cutoff to construct the corresponding counterfactual distributions.[8] Table 1 reports the estimated bunching magnitudes using alternative cutoff ages around 31. As the cutoff moves from the late twenties toward the early thirties, the estimated bunching becomes larger and more statistically significant until it reaches its peak at age 31. When the cutoff is shifted to age 32 afterwards, the estimated magnitude declines, and becomes less significant.

[8]For candidate cutoffs from age 28 to age 32, we re-estimate bunching using half-window widths from 1 to 5, chosen to align the left boundary of the bunching window with that in the main specification using age 31 as the cutoff. Since alternative cutoffs may not naturally satisfy the integration constraint, we adjust the counterfactual distribution following Chetty et al. (2011). See Section D for details.

Table 1: **Bunching Estimates at Alternative Cutoffs near Age 31**

Cutoff	Bunching Estimate ($\hat{B}$)	Standard Error
≤ 27	Unable to estimate	–
28	2.78	2.47
29	3.56*	2.17
30	7.17***	1.69
31	8.22***	1.34
32	6.10***	1.04
33	1.01	0.98
≥ 34	Unable to estimate	–

Notes: This table reports bunching estimates obtained by re-estimating the baseline bunching specification at alternative candidate cutoff ages near 31. The estimate $\hat{B}$ measures the normalized excess mass immediately to the left of each candidate cutoff, with standard errors reported in the third column. For candidate cutoffs from ages 28 to 32, the half-window h is allowed to vary from 1 to 5 to keep the lower boundary of the bunching window comparable across specifications; all other parameters are held fixed. Cutoffs at ages 27 or below and 34 or above cannot be estimated because they do not provide a suitable window for constructing a smooth counterfactual distribution.

For candidate cutoffs at age 27 or below and at age 34 or above, we do not report bunching estimates, as the smooth-counterfactual assumption cannot be implemented reliably near the boundaries of the relevant age range.[9] Moreover, women's strategic behavior is induced only by age cutoffs that are associated with salient age-related penalties. Ages far from 31 are unlikely to carry such penalties, and therefore do not generate excess mass around the corresponding cutoffs.

These results suggest that women's responses span a range of ages around the threshold rather than clustering at a single point. Nevertheless, the bunching estimate peaks sharply at age 31, supporting our use of age 31 as the representative norm threshold and justifying the single-cutoff abstraction in the theoretical framework. To further address the diffuse nature of the norm, we pool ages 30, 31, and 32 into a combined cutoff range and re-estimate the bunching response. Figure 3 plots the surgery-age distribution for this combined window: excess mass accumulates across ages 30–32 and falls off just beyond it, yielding an even larger bunching estimate. The figure shows that women's strategic retiming is concentrated in the early-thirties window, consistent with a norm whose salience spans this narrow band.

[9] The bunching estimator recovers a smooth counterfactual from the distribution outside the excluded window around the cutoff, which requires sufficient support on both sides. At cutoffs of 27 or below (34 or above), the lower (upper) side offers too few adjacent age bins, so the counterfactual is driven by extrapolation rather than local smoothness, making excess-mass estimates unreliable and not comparable to those for ages 28–33.

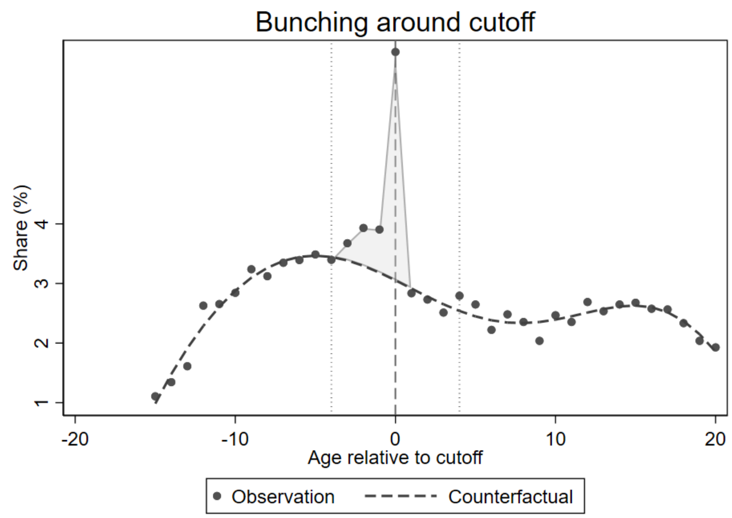


Figure 3: Bunching Estimation combining Cutoffs 30, 31, 32

Notes: This figure reports the bunching estimate when the cutoff is defined more broadly to combine ages 30, 31, and 32. This specification allows the age-based norm to operate over a wider age range rather than only at the baseline age-31 cutoff. The exercise assesses whether excess mass is concentrated near the broader threshold region, consistent with a more diffuse perception of the "leftover women" age norm.

5 Surgery Types, Marriage Pressure, and Economic Opportunities

To understand the social and economic conditions under which the strategic response is stronger or weaker, we examine heterogeneity along four dimensions: surgery type, the intensity of fertility norms, marriage-market competition, and women's labor-market opportunities.

5.1 Surgery Types

We first ask whether the bunching pattern varies systematically across procedure types. Facial features and breast appearance are well-documented primary dimensions of female attractiveness in partner evaluation. If the bunching is driven by marriage-market incentives, procedures targeting these body parts should exhibit more pronounced pre-cutoff excess mass than procedures with lower marriage-specific salience.

Panel 1 of Table 2 (Figure A6) separates cosmetic admissions into two groups: face and breast procedures, and all other cosmetic procedures. Bunching is substantially larger for face and breast surgeries, with an estimated coefficient of $\hat{B} = 9.99$ (SE = 1.34), than for other cosmetic procedures ($\hat{B} = 6.47$ (SE = 1.85)). Women selectively retime the procedures with the highest anticipated marital returns. Having established this within-woman pattern,

Table 2: **Heterogeneity Analysis Results**

Category	Measure	Subsample	$\hat{B}$	N
Procedure type	Procedure category	Face and breast	9.99*** (1.34)	13,992
		Other body regions	6.47*** (1.85)	18,033
Social norm	Fertility norm	Strong	14.48*** (2.25)	12,599
		Weak	4.96*** (1.48)	19,426
	Confucian culture	High	8.03*** (2.15)	8,163
		Low	8.27*** (1.46)	23,862
Marriage pressure	Female-to-male ratio	High ratio	9.49*** (1.66)	16,177
		Low ratio	6.69*** (1.72)	15,848
	Baidu search index	High search intensity	9.26*** (1.48)	18,757
		Low search intensity	6.56*** (2.18)	13,268
Economic opportunities	Economic dependence	High economic dependence	13.66*** (2.20)	19,453
		Low economic dependence	2.80** (1.31)	12,572
	Educational attainment	Low female education	13.12*** (2.37)	11,773
		High female education	5.83*** (1.12)	20,252
	Gender wage gap	High wage gap	10.09*** (1.70)	21,909
		Low wage gap	5.00*** (1.52)	10,116

Notes: This table summarizes the heterogeneity analysis. Column 1 reports the broad heterogeneity dimension, Column 2 the measure used to operationalize it, and Column 3 the subsamples defined under each measure. Column 4 reports the estimated bunching magnitude $\hat{B}$, with standard errors in parentheses, and Column 5 the number of observations in each subsample. Fertility norm intensity is based on the 2010 CFPS question on the importance of having children to carry on the family line; provinces are classified as strong or weak relative to the national average. Confucian culture is measured by prefectural clan density following Greif and Tabellini (2017), with provinces classified as high or low relative to the national mean. The Baihe.com measure is the female-to-male ratio of registered active users in a city; the Baidu search index is constructed from city-level search intensity for marriage-pressure keywords. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

we turn to cross-regional variation: does the response intensify where the marriage-market pressure itself is stronger?

5.2 Fertility Norm

Our central interpretation is that the bunching reflects a response to a fertility-driven age deadline in the marriage market. If so, it should be most pronounced precisely where fertility norms are strongest. We test this prediction by examining heterogeneity along this dimension.

Fertility willingness. To capture fertility-related family norms, we use the 2010 CFPS question, "It is important to have children to carry on the family line," answered on a 1–5 scale from "strongly disagree" to "strongly agree," and we classify provinces as having strong or weak fertility norms relative to the national average. Panel 2 of Table 2 (Figure A7) reports the corresponding surgery-age distributions. Bunching is far more pronounced in provinces with stronger fertility norms, $\hat{B}$ = 14.48 (SE = 2.25), than in those with weaker norms, $\hat{B}$ = 4.96 (SE = 1.48). This gradient is precisely what the "leftover women" mechanism predicts: where continuing the family line is more highly valued, the marriage-market penalty attached to the perceived fertility deadline is sharper, giving women stronger incentives to invest in appearance before the age at which their marriage and childbearing prospects are seen to decline.

Confucian culture. Because fertility norms are strongly correlated with broader patriarchal values, the preceding result admits an alternative interpretation: stronger bunching may reflect not a response to a fertility *deadline*, but merely that women in more patriarchal places conform more readily to marriage-market pressures in general. To separate these observationally similar explanations, we exploit a proxy for patriarchal values that is strong where gender norms are conservative yet carries *no* built-in age threshold: Confucian culture.[10] The generalized-patriarchy explanation thus predicts stronger bunching in more Confucian areas, whereas our age-deadline mechanism predicts none. Following Greif and Tabellini (2017), we measure the intensity of Confucian culture using prefectural clan density and classify provinces as above or below the national mean. Panel 2 of Table 2 (Figure A8) reveals no meaningful heterogeneity: the bunching coefficient is virtually identical in high- and low-Confucian-culture provinces ($\hat{B}$ = 8.03, SE = 2.15 versus $\hat{B}$ = 8.27, SE = 1.46). The data therefore favor the age-deadline interpretation over the generalized-patriarchy alternative: the strategic beauty investment we document is driven by an age-specific marriage-market penalty, not by conservative gender norms per se.

5.3 Marriage Pressure

We next test whether the bunching response intensifies where marriage-market competition is more intense. Where women face a tighter local marriage market, crossing the age threshold unmarried carries a greater cost, which strengthens the incentive to invest in appearance before the cutoff.

Marriage-market competition. We then measure competition by the ratio of active female to male users on Baihe.com, one of China's largest matchmaking platforms; a higher ratio means women compete for a smaller pool of partners. Classifying provinces as high-

[10]Confucianism is the foundational framework for traditional gender norms in East Asia (Slote and De Vos, 1998) and prescribes a clearly subordinate role for women, yet it attaches no penalty to any particular age. If anything, a woman's authority within the traditional Chinese family *rises* with age as she moves from daughter-in-law to mother and eventually to mother-in-law (Baker, 1979).

or low-competition relative to the national mean, Panel 3 of Table 2 (Figure A9) compares their surgery-age distributions. Bunching is markedly stronger in high-competition provinces, $\hat{B} = 9.49$ (SE = 1.66), than in low-competition ones, $\hat{B} = 6.69$ (SE = 1.72). The contrast fits the "leftover women" penalty as the primary force: where competition for partners is fiercer, the incentive to invest in appearance before the threshold is stronger.

Perceived marriage pressure. We next construct a complementary, perception-based proxy from the Baidu Search Index. Provided by China's dominant search engine, the index is the Chinese counterpart to Google Trends: it records how intensively users in a given location search for a term, and thus tracks the local salience of an issue over time. We collect city-level search volumes for marriage-pressure keywords such as "maternal age" and "urged marriage," averaged over 2017–2019, and classify cities as high- or low-intensity according to whether their search frequency lies above or below the national mean. Because people search for these terms when marriage pressure is personally salient, the index offers a behavioral, demand-side measure of felt urgency that complements the supply-side competition proxy above. The pattern mirrors the competition result: Panel 3 of Table 2 (Figure A10) reports a bunching coefficient of $\hat{B} = 9.26$ (SE = 1.48) in high-intensity cities, compared with $\hat{B} = 6.56$ (SE = 2.18) in low-intensity cities.

All these measures point in the same direction: the pre-threshold concentration of beauty investment is sharper where marriage-market pressure is more intense, whether measured by objective competition or by subjectively perceived urgency.

5.4 Women's Economic Opportunity

We next turn to women's labor-market opportunities, which give them an alternative to appearance for improving their marriage-market prospects. While appearance investment enhances perceived attractiveness, women with stronger labor-market prospects can improve their marriage-market standing through additional channels—such as earnings potential and occupational status—that partially substitute for appearance-based signals. The relative return to beauty investment should therefore be lower when these alternative margins are more valuable, and we correspondingly expect bunching to be attenuated among women with stronger economic opportunities. Panel 4 of Table 2 tests this prediction by examining heterogeneity along three dimensions: economic dependence, educational attainment, and local gender wage gaps.

Economic Dependence. Using patient-level occupational information, we classify women as having high economic dependence if they are students, unemployed, or engaged in informal or unofficial jobs, and as having *low economic dependence* if they hold formal positions, such as workers, government or institutional staff, and company employees. This classification captures the stability of a woman's labor market engagement rather than her level of earnings. Figure A11 shows that the bunching pattern is substantially stronger among women with high economic dependence. For this group, the estimated coefficient is $\hat{B} = 13.66$ (SE = 2.20), compared with $\hat{B} = 2.80$ (SE = 1.31) among women with low economic dependence. This divergence is consistent with the proposed mechanism: women with weaker labor market opportunities face a higher relative return to marriage-market investment, which strengthens the incentive to undergo surgery before the age threshold.

Female Educational Attainment. A second proxy for women's economic opportunities is regional female educational attainment. Where women are better educated, they have stronger labor-market attachment and higher earnings potential, reducing the relative return to beauty investment as a matrimonial strategy. We therefore expect bunching to be weaker in high-attainment provinces. Using the 2015 population census, we measure female educational attainment as the average educational level of women aged 25–35 in each province and classify provinces as above or below the national average. Figure A12 reveals striking heterogeneity: in low-attainment provinces, the bunching coefficient is $\hat{B} = 13.12$ (SE = 2.37), while in high-attainment provinces it falls by more than half, to $\hat{B} = 5.83$ (SE = 1.12). The pattern confirms that beauty investment is most valuable as a matrimonial strategy where women's economic alternatives are most constrained.

Gender Wage Gap. A third proxy captures structural disadvantage in the labor market. Where gender wage gaps are wider, women face lower returns to labor-market participation relative to men, weakening their social-economic status and raising the relative value of securing a favorable match through appearance. We therefore expect bunching to be stronger in high-gap provinces. Following Blau and Kahn (2017), we define the gender wage gap as the difference between male and female weighted average wages across industries, with weights reflecting gender-specific employment shares within each province; this measure captures the contribution of occupational segregation to women's labor-market disadvantage independent of educational attainment (construction details in Appendix E). Figure A13 confirms the prediction: the bunching coefficient in high-gap provinces is $\hat{B} = 10.09$ (SE = 1.70), compared with $\hat{B} = 5.00$ (SE = 1.52) in low-gap provinces.

Taken together, the outside-option results deliver a consistent message: beauty investment is most valuable as a matrimonial strategy where women's alternatives outside marriage are most constrained. Section 6 develops a theoretical model that rationalizes all three dimensions of heterogeneity through a unified mechanism.

5.5 Alternative Explanations

Having shown that the bunching response strengthens with marriage-market incentives, we now consider four alternative interpretations: that the pre-cutoff concentration reflects medical considerations, budget constraints, career incentives, or postpartum recovery demand, rather than the marriage-market mechanism we propose. Below, we provide evidence inconsistent with each of these alternative explanations.

Medical considerations. A natural alternative is that women undergo these procedures before their early thirties for medical reasons unrelated to marriage—for instance, certain procedures may be medically recommended to be completed before a given age. Yet existing medical guidance offers little support for this account, particularly for the two procedure types driving our result—face and breast surgeries (Section 5.1). Breast augmentation is approved from the early twenties and routinely performed well into a woman's mid-thirties (Jordan and Corcoran, 2013), while facial rejuvenation is typically recommended even later, once visible signs of aging appear in the forties (Meretsky et al., 2024). A related concern is that surgery grows medically riskier with age—healing slows and anesthesia becomes harder to tolerate—prompting women to operate while still young (Luca et al., 2023). But if medical necessity were driving the pattern, bunching should be similarly pronounced across all regions. Instead, our heterogeneity analysis shows that bunching strength varies systematically across regions and is closely tied to marriage-market incentives (Section 5.3). This evidence weighs against the medical explanation and instead supports the marriage-market interpretation.

Budget constraints. Another alternative is that the pre-cutoff concentration reflects a gradually relaxing budget constraint rather than a marriage-market deadline: cosmetic surgery is costly, and women may simply need until their late twenties to accumulate sufficient resources to afford it. Under this account, the pre-31 spike would mark the point at which affordability catches up with demand, rather than a strategic response to an approaching age penalty. Conceptually, however, this explanation does not naturally predict bunching at a common age. Because women differ in income and savings, the age at which any given woman first becomes able to afford surgery should itself be heterogeneous, generating a gradually rising distribution of surgery timing rather than a sharp spike concentrated at a single age. The marriage-market norm, by contrast, imposes the same age threshold on all women, which is precisely what generates a common bunching point. Empirically, we find no evidence of a relaxing budget constraint around age 31: observable income evolves smoothly through this age (Figure A3), and two additional financial measures reinforce this conclusion. Household savings and intergenerational transfers—an alternative channel through which women might finance surgery, for instance from parents—both evolve smoothly through the threshold, with no discernible acceleration as women approach age 31 (Figure A3).

Career incentives. It is also plausible that the bunching reflects labor-market rather than marriage-market incentives. Given the well-documented beauty premium in hiring and wages (Hamermesh, 2011), women may advance cosmetic surgery to improve their career prospects as they approach their peak working years. To assess this possibility, we ask whether local cosmetic-surgery demand co-moves with marriage-related or with employment-related concerns.

We measure these concerns with the Baidu Search Index introduced in Section 5.3. From the search intensity of marriage-related terms—"maternal age" and "urged marriage"—we build a city-by-year index of marriage pressure, combining the terms into an aggregate index by principal component analysis,[11] and we construct an analogous index of labor-market pressure from employment-related terms, "recruitment" and "employment."

We regress the log number of cosmetic procedures performed in a city-year on each index, controlling for city and year fixed effects (Appendix Table A3). Cosmetic-surgery demand rises with marriage-related pressure: the association is positive and statistically significant, for both the aggregate index and the individual terms. By contrast, it bears no such relationship to employment-related pressure, whose coefficients are small and statistically insignificant. This asymmetry suggests that cosmetic-surgery demand is driven more by marriage-market than by labor-market concerns, consistent with the marriage-market interpretation of the bunching.

Post-childbirth recovery. A final alternative is that the bunching reflects cosmetic demand following childbirth, as postpartum women seek to reverse the physical changes of pregnancy.[12] Because age 31 lies near peak childbearing years, the surgery spike might simply track a concentration of births. If so, childbirth timing should itself bunch at the same

[11] To reduce idiosyncratic noise in individual keywords, we also construct an aggregate index that combines the constituent terms into a single composite, defined as the first principal component of the standardized search-intensity series.

[12] Procedures that bundle breast and abdominal surgery to restore the pre-pregnancy body, marketed as the "mommy makeover," have become one of the most sought-after cosmetic-surgery combinations, with demand rising steadily over the past two decades. In China, a parallel post-childbirth recovery (*chanhou xiufu*) industry has expanded rapidly, with the postpartum care and recovery market reaching RMB 59.4 billion in 2023. See `https://www1.hkexnews.hk/listedco/listconews/sehk/2025/0626/11726058/sehk24123102248.pdf`.

age. It does not: in the 2015 one-percent Population Census, the age distribution of mothers at first birth evolves smoothly through age 31, with no kink or spike (Figure A14), and the number of children rises smoothly across the cutoff in the 2021 CGSS (Figure A3). Because fertility timing is smooth precisely where cosmetic surgery bunches, post-childbirth demand is unlikely to account for the observed pattern.[13]

6 A Search Model with Age-based Marriage Norm

The empirical analysis has established a consistent set of patterns: cosmetic surgery bunches sharply just before the age threshold, which is more pronounced for marriage-relevant procedures and only among women, with the response intensifying where marriage-market stakes are higher and women's economic opportunities are weaker. We now show that a single mechanism can organize all of them. We develop a continuous-time search model in which a woman chooses when to undertake a one-time, irreversible beauty investment that raises her matching rate. We assume that a "leftover women" social norm penalizes those who remain unmarried past a salient age. We show that the norm acts as a deadline: it pulls women whose ideal surgery timing lies just beyond the threshold back to it, generating the observed bunching, while its comparative statics map onto the documented heterogeneity. This section presents the setup, conveys the key intuitions, and states the model's predictions; the complete formal development, derivations, and proofs are relegated to Appendix F.

Setup. We consider a population of single women. A woman's marital appeal has two components: her *beauty*, or physical attractiveness, and her *non-beauty appeal*, defined as everything else the marriage market rewards, such as education, earnings, and occupational standing. Each woman chooses when, if ever, to undergo beauty-enhancement surgery, a binary, one-time, and irreversible decision that raises her beauty and hence the rate at which she is matched. Surgery carries a fixed cost, but she incurs it only if she is still single at her chosen age; a woman who marries beforehand never pays. In addition, the marriage market is governed by a social norm that stigmatizes women who remain unmarried past a culturally salient age threshold, beyond which their matching prospects decline sharply.

A match yields a fixed lifetime value, and each woman chooses her surgery timing to maximize expected discounted lifetime utility. Two assumptions discipline the model. First,

[13] As a more direct check, we draw on fertility records in our inpatient database to identify women who gave birth during the sample period prior to undergoing surgery. Matching on patient name, admission date, and hospital, we identify 688 such women—2% of the full female cosmetic surgery sample and 6.5% of female patients aged 27–31 over the study period. Our main results are robust to excluding these women from the sample (Figure A15).

beauty and non-beauty appeal are substitutes: a higher non-beauty type implies stronger baseline marital prospects but a smaller marginal return to surgery. Second, the norm penalty applies to women of all types yet is heterogeneous in its incidence—it is milder for those with stronger non-beauty appeal, whose other attributes partially offset the forgone beauty premium.

Optimal timing balances the marginal benefit of advancing surgery against its marginal cost. The benefit is the gain in matching probability: surgery raises a woman's matching prospect. The cost has two components: she incurs the expenditure earlier, forgoing the return on those resources in the interim, and she bears the risk that the procedure proves unnecessary—had she waited, she might have matched naturally and avoided the cost altogether. A woman undergoes surgery at the date where the benefit just exceeds the cost of waiting one period longer.

Bunching at the norm threshold. Absent the norm, these choices aggregate across women with heterogeneous non-beauty attributes into a smooth distribution of surgery ages. The norm breaks this smoothness. Because crossing the threshold reduces the matching return to surgery, a woman's lifetime payoff falls discretely at the cutoff. The norm raises the marginal benefit of operating early—surgery now secures both the matching boost and the return that the penalty would otherwise erode—while leaving the marginal cost unchanged, tilting the trade-off toward earlier timing.

The threshold thus acts as a one-sided "magnet." Consider a woman whose unconstrained optimum lies just above the cutoff. She gains almost nothing from delaying to that point, since her payoff is nearly flat over this range; yet delaying pushes the procedure across the threshold, forfeiting a discrete share of its return. She is therefore better off advancing the surgery to the cutoff, capturing essentially the same timing benefit while escaping the penalty. A woman whose optimum already lies below the cutoff faces no such force to advance, because her payoff declines smoothly as she approaches it. Relocation therefore flows in one direction only—women advance inward from above, never from below.

The model thus predicts a mass point in the surgery-age distribution exactly at the threshold, formed by women drawn down from just above it. This one-sided bunching matches the pattern documented in Section 4.3: a visible spike at the threshold age and no offsetting deficit on the left. Appendix F.4 establishes the result formally.

Comparative statics. The comparative statics delivers four sharp predictions about when bunching is stronger, each mapping to a distinct dimension of heterogeneity in the data. Appendix F.6 shows them formally as corollaries; we present their intuitions in this section and connect each to its empirical counterpart.

The first concerns the salience of beauty. When beauty carries more weight in partner evaluation, surgery delivers a larger boost, so the norm's multiplicative penalty strips away a larger absolute amount at the threshold. The payoff drop widens and more women relocate, implying that bunching is sharper for procedures whose beauty content is most marriage-relevant. This is the contrast documented in Section 5.1: face and breast surgeries, the features weighted most heavily in partner evaluation, bunch more sharply than other cosmetic procedures.

The second concerns how thin the local marriage market is. Where partners are scarce, two forces both sharpen bunching: matches arrive slowly, so the boost from surgery is worth more and the penalty for crossing the threshold is larger. It implies that bunching is stronger where women face a thinner pool of prospective partners. Section 5.3 bears this out: bunching is more pronounced in high-pressure provinces, where female-to-male ratios on matchmaking platforms are highest.

The third concerns women's non-beauty appeal—their marriage-market standing apart from looks. Where this appeal is stronger, two forces weaken bunching: women already match more readily, leaving surgery less marital surplus to add, and beauty matters less at the margin for those already attractive on other dimensions. This matches the outside-option results in Section 5.4: bunching concentrates among women with high economic dependence, lower educational attainment, and wider local gender wage gaps.

Finally, the mechanism rests on two ingredients—a marriage-market return to beauty and a gendered age penalty—and removing either one shuts bunching down. No bunching arises for procedures without marriage-market content, nor among men, who face no analogous age norm. These two placebos, borne out by the distributions for functional surgeries and for men in Section 4.3, confirm that neither ingredient alone suffices.

7 Conclusion

This paper studies whether and how women strategically invest in their physical appearance in response to marriage-market incentives. We exploit China's "leftover women" norm as an exogenously determined source of variation in the returns to beauty investment. Drawing on a nationwide inpatient hospital dataset, we document substantial bunching in cosmetic surgery admissions immediately before the early thirties. This excess mass is specific to aesthetic procedures, absent for functional surgeries and for men, and is more pronounced in regions with higher marriage-market competition, and weaker economic opportunities for women.

The findings carry several implications for our understanding of gender inequality. First, the costs of the "leftover women" norm are borne asymmetrically across genders: men face no analogous age-specific penalty and exhibit no strategic surgery response at any age. The financial expenditure, surgical risk, and recovery time entailed by pre-cutoff beauty investment thus constitute an implicit "pink tax"—a welfare cost falling entirely on women. Second, social norms can generate significant distortions in costly, irreversible investment decisions. The "leftover women" norm introduces an artificial deadline, inducing women to advance procedures earlier than individually optimal. This distortion falls most heavily on women with the fewest economic alternatives, for whom beauty investment serves as a substitute for human capital in accessing the marriage market, potentially reinforcing existing inequalities. Third, our results suggest that policies aimed at reducing gender wage gaps and expanding female educational opportunities carry benefits beyond their direct labor-market effects: by strengthening women's social-economic status, they attenuate the behavioral distortions induced by age-based marriage-market norms.

References

Adda, J., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy 125*(2), 293–337.

Ashraf, N., N. Bau, N. Nunn, and A. Voena (2020). Bride price and female education. *Journal of Political Economy 128*(2), 591–641.

Baker, H. D. (1979). *Chinese family and kinship*. Columbia University Press.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender identity and relative income within households. *Quarterly Journal of Economics 130*(2), 571–614.

Bhaskar, V. and E. Hopkins (2016). Marriage as a rat race: Noisy premarital investments with assortative matching. *Journal of Political Economy 124*(4), 992–1045.

Bhaskar, V., W. Li, and J. Yi (2023). Multidimensional premarital investments with imperfect commitment. *Journal of Political Economy 131*(10), 2893–2919.

Blau, F. D. and L. M. Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature 55*(3), 789–865.

Broekmans, F. J., E. A. Knauff, E. R. te Velde, N. S. Macklon, and B. C. Fauser (2007). Female reproductive ageing: current knowledge and future trends. *Trends in Endocrinology & Metabolism 18*(2), 58–65.

Bursztyn, L., T. Fujiwara, and A. Pallais (2017). "acting wife": Marriage market incentives and labor market investments. *American Economic Review 107*(11), 3288–3319.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review 110*(10), 2997–3029.

Bursztyn, L. and D. Y. Yang (2022). Misperceptions about others. *Annual Review of Economics 14*(1), 425–452.

Chen, Z., Z. Liu, J. C. Suárez Serrato, and D. Y. Xu (2021). Notching r&d investment with corporate income tax cuts in china. *American Economic Review 111*(7), 2065–2100.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The quarterly journal of economics 126*(2), 749–804.

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2009). Investment in schooling and the marriage market. *American Economic Review 99*(5), 1689–1713.

Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque (2012). Fatter attraction: Anthropometric and socioeconomic matching on the marriage market. *Journal of Political Economy 120*(4), 659–695.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review 107*(8), 2109–2167.

Dee, T. S., W. Dobbie, B. A. Jacob, and J. Rockoff (2019). The causes and consequences of test score manipulation: Evidence from the new york regents examinations. *American Economic Journal: Applied Economics 11*(3), 382–423.

Dunson, D. B., D. D. Baird, and B. Colombo (2004). Increased infertility with age in men and women. *Obstetrics & Gynecology 103*(1), 51–56.

Dunson, D. B., B. Colombo, and D. D. Baird (2002). Changes with age in the level and duration of fertility in the menstrual cycle. *Human Reproduction 17*(5), 1399–1403.

Fincher, L. H. (2023). *Leftover Women: The Resurgence of Gender Inequality in China* (10th Anniversary ed.). London: Bloomsbury Academic.

Fisman, R., S. S. Iyengar, E. Kamenica, and I. Simonson (2006). Gender differences in mate selection: Evidence from a speed dating experiment. *The Quarterly Journal of Economics 121*(2), 673–697.

Gaetano, A. (2014). "leftover women": Postponing marriage and renegotiating womanhood in urban china. *Journal of Research in Gender Studies 4*(2), 124–149.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Graham, J. R., C. R. Harvey, and M. Puri (2017). A corporate beauty contest. *Management Science 63*(9), 3044–3056.

Graham, N. (2015). Leftover women: The resurgence of gender inequality in china.

Grand View Research (2024). Cosmetic surgery and procedure market size, share & trends analysis report by procedure type, by gender, by age group, by region, and segment forecasts, 2025–2033. Technical report, Grand View Research. Accessed: April 2026.

Greif, A. and G. Tabellini (2017). The clan and the corporation: Sustaining cooperation in china and europe. *Journal of Comparative Economics 45*(1), 1–35.

Hamermesh, D. S. (2011). *Beauty pays: Why attractive people are more successful*. Princeton University Press.

Hamermesh, D. S. and J. E. Biddle (1994). Beauty and the labor market. *American Economic Review 84*(5), 1174–1194.

Han, L. and X. Shi (2019). How does intergenerational investment respond to changes in the marriage market? evidence from china. *Journal of Development Economics 141*, 102310.

Han, L., X. Shi, and M.-a. Zhang (2022). How does matching uncertainty affect marital surplus? theory and evidence from china. Working paper, HKUST Center for Economic Policy.

Han, L., X. Shi, and M.-a. Zhang (2026). Ignorance is whose bliss: The repeal of compulsory premarital health examinations and marital outcomes in rural china. *Journal of the European Economic Association*, jvag026.

Hitsch, G. J., A. Hortaçsu, and D. Ariely (2010). Matching and sorting in online dating. *American Economic Review 100*(1), 130–163.

ISAPS (2023, June). Survey finds surgical aesthetic procedures on the rise globally. `https://www.isaps.org/articles/isaps-blog/2024/isaps-2023-global-survey-results/`. Accessed: 2026-04-10.

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review 69*(3), 576–595.

Ji, Y. (2015a). Between tradition and modernity: "leftover" women in shanghai. *Journal of Marriage and Family 77*(5), 1057–1073.

Ji, Y. (2015b). Between tradition and modernity: "leftover" women in shanghai. *Journal of Marriage and Family 77*(5), 1057–1073.

Jordan, S. W. and J. F. Corcoran (2013). Considerations in breast augmentation in the adolescent patient. *Seminars in Plastic Surgery 27*(1), 67–71.

Kleven, H., C. Landais, and G. Leite-Mariante (2025). The child penalty atlas. *Review of Economic Studies 92*(5), 3174–3207.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Kuhn, P. and K. Shen (2013). Gender discrimination in job ads: Evidence from China. *Quarterly Journal of Economics 128*(1), 287–336.

Liu, Q. (2021). Relational dignity, state law, and chinese leftover women's choices in marriage and childbearing. *Asian Journal of Law and Society 8*(1), 151–167.

Low, C. (2024a). The human capital–reproductive capital trade-off in marriage market matching. *Journal of Political Economy 132*(2), 552–576.

Low, C. (2024b). Pricing the biological clock: The marriage market costs of aging to women. *Journal of Labor Economics 42*(2), 395–426.

Luca, E., C. Schipa, C. Cambise, L. Sollazzi, and P. Aceto (2023). Implication of age-related changes on anesthesia management. *Saudi Journal of Anaesthesia 17*(4), 474–481.

Malhotra, A. and S. Elnakib (2021). Evolution in the evidence base on child marriage 2000–2019. Prepared on behalf of the UNFPA–UNICEF Global Programme to End Child Marriage.

Meretsky, C. R., P. Hausner, B. P. Flynn, and A. T. Schiuma (2024). A systematic review and comparative analysis of reconstructive rhytidectomy: advantages, disadvantages, and patient outcomes. *Cureus 16*(10), e71006.

Mobius, M. M. and T. S. Rosenblat (2006). Why beauty matters. *American Economic Review 96*(1), 222–235.

Nanda, P., R. K. Verma, and J. Abrahamson (2012). Child marriage in southern asia: Context, evidence and policy options for action. Developed by ICRW as commissioned by AFPPD; supported by AusAID.

Navot, D., R. Bergh, M. A. Williams, G. J. Garrisi, I. Guzman, B. Sandler, and L. Grunfeld (1991). Poor oocyte quality rather than implantation failure as a cause of age-related decline in female fertility. *The Lancet 337*(8754), 1375–1377.

Peters, M. and A. Siow (2002). Competing premarital investments. *Journal of Political Economy 110*(3), 592–608.

Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector. *Labour Economics 14*(3), 371–391.

Porter, J. and P. Yu (2015). Regression discontinuity designs with unknown discontinuity points: Testing and estimation. *Journal of Econometrics 189*(1), 132–147.

Slote, W. H. and G. A. De Vos (1998). *Confucianism and the Family.* State University of New York Press.

To, S. (2013). Understanding sheng nu ("leftover women"): The phenomenon of late marriage among chinese professional women. *Symbolic Interaction 36*(1), 1–20.

Wei, S.-J. and X. Zhang (2011). The competitive saving motive: Evidence from rising sex ratios and savings rates in china. *Journal of political Economy 119*(3), 511–564.

Wu, Y., S. Mulkens, and J. M. Alleva (2022). Body image and acceptance of cosmetic surgery in china and the netherlands: A qualitative study on cultural differences and similarities. *Body Image 40*, 30–49.

You, J., X. Yi, and M. Chen (2021). Love, life, and "leftover ladies" in urban china: Staying modernly single in patriarchal traditions. *China Economic Review 68*, 101626.

Zamberlan, A. and P. Barbieri (2023). A 'potential motherhood' penalty? a longitudinal analysis of the wage gap based on potential fertility in Germany and the United Kingdom. *European Sociological Review 39*(6), 920–934.

Zhang, J., S. Fei, and Y. Wen (2023). How does the beauty of wives affect post-marriage family outcomes? Helen's face in Chinese households. *Journal of Economic Behavior & Organization 212*, 122–137.

Supplemental Appendix

A Supplemental Figures and Tables

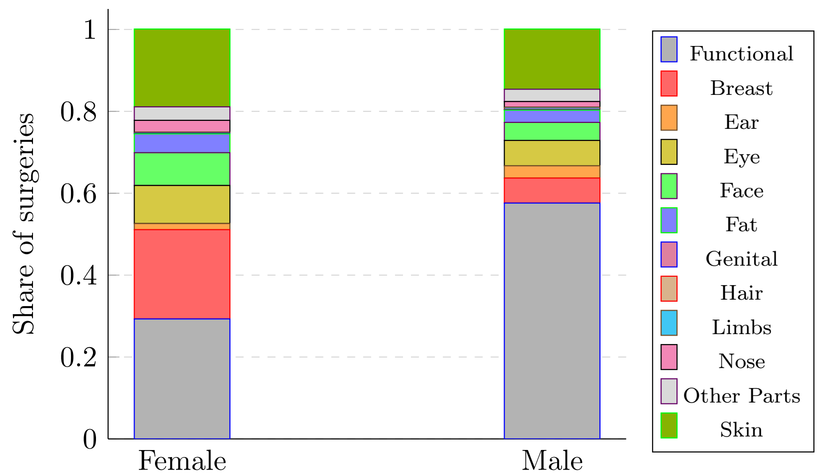
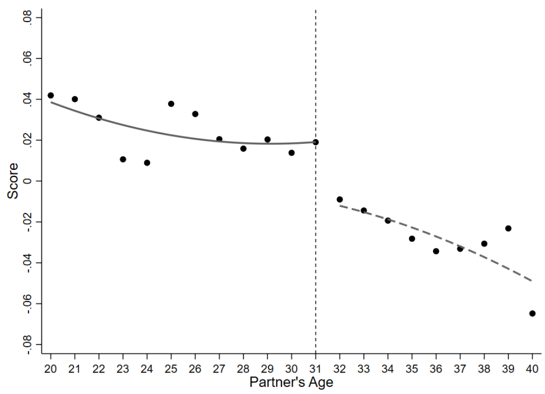
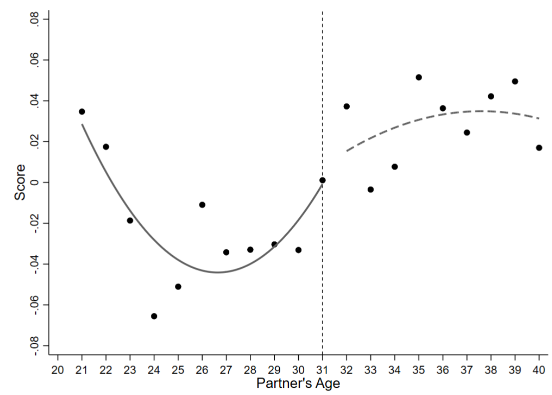


Figure A1: Distribution of Surgery Types by Gender

Notes: This figure shows the distribution of surgery types by patient gender. Each stacked bar reports the share of surgeries accounted for by functional procedures and by each non-functional body-part category. Shares are computed separately for female and male patients and sum to one within gender.



Men rate women, $c = -0.019^{**}(0.0078)$



Women rate men, $c = 0.010(0.034)$

Figure A2: Perceived Attractiveness around age 31

Notes: This figure uses data from the 2021 Chinese General Social Survey (CGSS) to illustrate how perceived attractiveness varies with partners' ages. The sample includes male and female respondents who participated in three rounds of a randomized experiment. In each round, respondents were shown randomized information about a potential partner across six dimensions and were then asked to rate that person's marital attractiveness. The x-axis represents the randomized age of the potential partner, and the y-axis shows the residual from regressing the attractiveness score on the other attributes of the potential partner, controlling for respondents' characteristics and survey round fixed effects. The coefficient c reported above each panel measures the estimated discrete drop in the residualized attractiveness rating at the age-31 cutoff, with standard errors in parentheses. See C for the detailed sample construction and methodology.

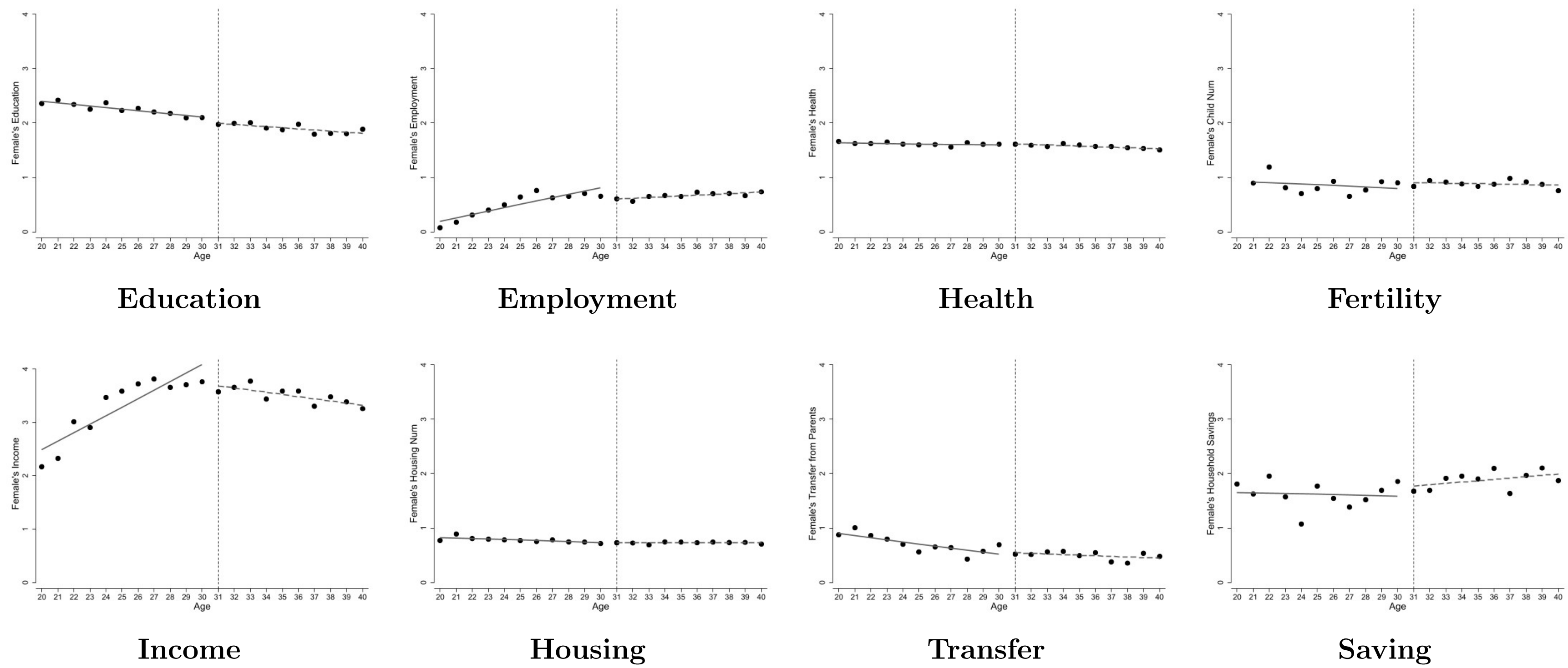
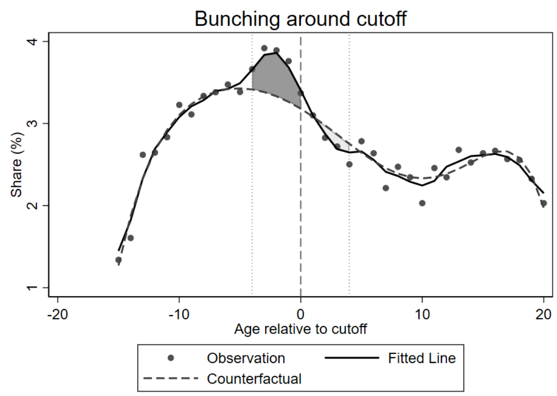
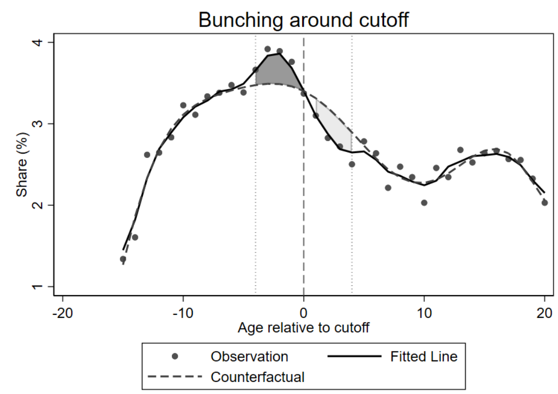


Figure A3: Observable Characteristics around the Age-31 Cutoff

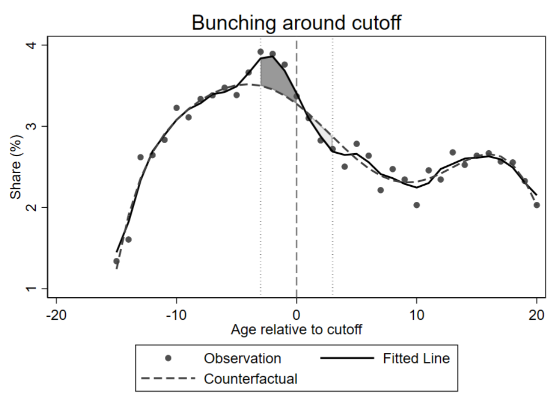
Notes: This figure plots the distribution of observable characteristics around age 31, including education, employment, health, fertility, income, housing, parental transfers, and household savings. These plots are used to assess whether there are discontinuous changes in predetermined or contemporaneous observables at the age-31 cutoff. The first six panels are constructed from the CGSS data. Since the CGSS does not contain detailed information on private transfers and savings, the last two panels are supplemented with data from the CFPS. The absence of comparable discontinuities in these variables supports the interpretation that the bunching pattern is specific to cosmetic surgery timing rather than driven by broader changes in sample composition around age 31.



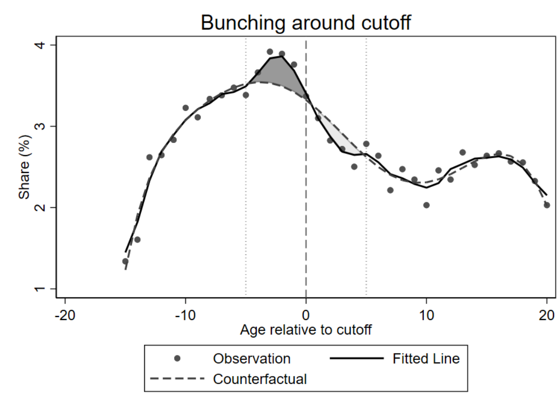
$Q = 6,\ \hat{B} = 6.88^{***}(1.25)$



$Q = 8,\ \hat{B} = 7.45^{***}(1.24)$



$h = 3,\ \hat{B} = 7.80^{***}(1.22)$



$h = 5,\ \hat{B} = 5.51^{***}(1.13)$

Figure A4: Robustness Checks: Alternative Parameter Choices

Notes: This figure reports robustness checks for the bunching estimates under alternative choices of estimator parameters. The top panels vary the polynomial order Q, and the bottom panels vary the excluded window h around the age-31 cutoff. In each panel, the horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.

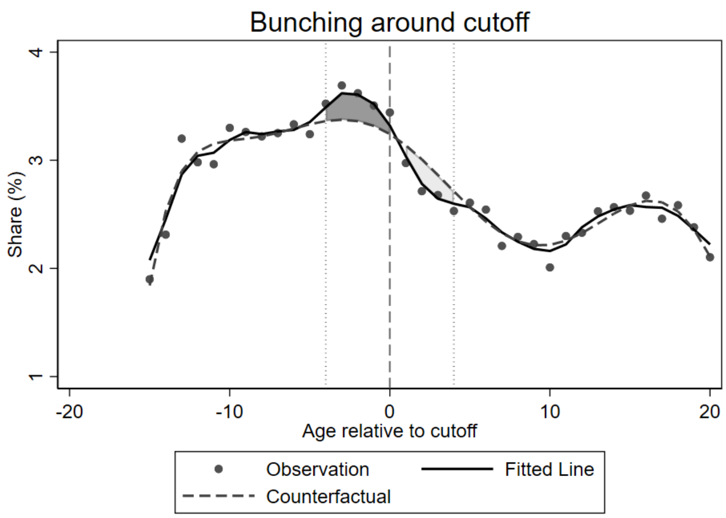
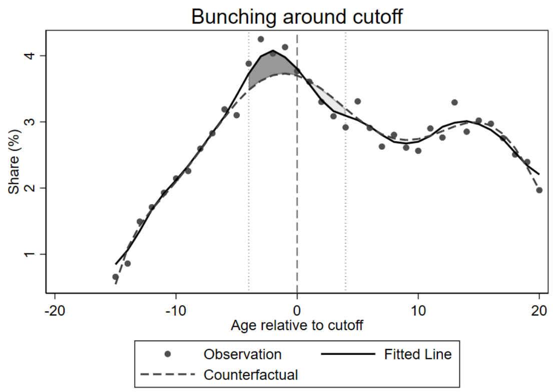


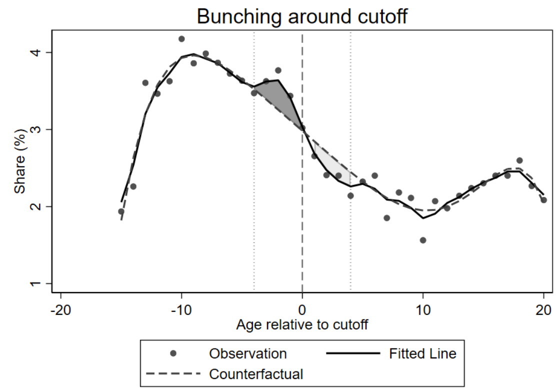
Figure A5: Robustness Checks: All Surgeries

$\hat{B} = 6.75^{***}(1.71)$

Notes: This figure plots female patients' age-at-first-surgery distribution around the age-31 cutoff, using all surgery types. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.



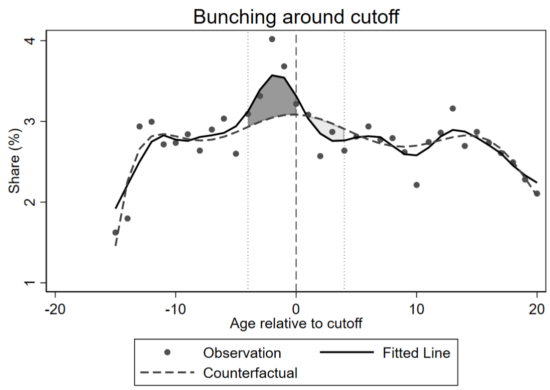
Face and Breast, $\hat{B} = 9.99^{***}(1.34)$



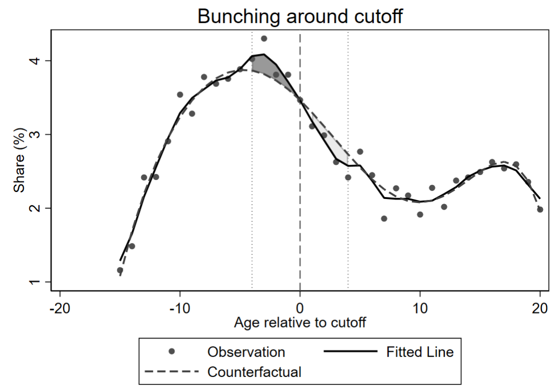
Other Types, $\hat{B} = 6.47^{***}(1.85)$

Figure A6: The Female Age Distribution by Cosmetic Surgery Types

Notes: This figure compares female patients' age-at-first-surgery distributions around the age-31 cutoff for face and breast cosmetic procedures and for other cosmetic procedures. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.



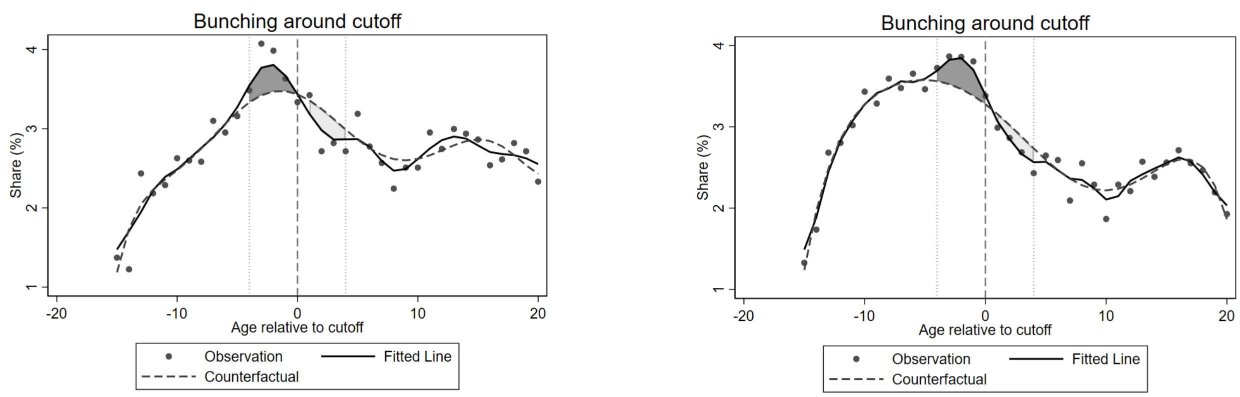
Strong Norm, $\hat{B} = 14.48^{***}(2.25)$



Weak Norm, $\hat{B} = 4.96^{***}(1.48)$

Figure A7: Female Age Distribution by Fertility Norm Intensity

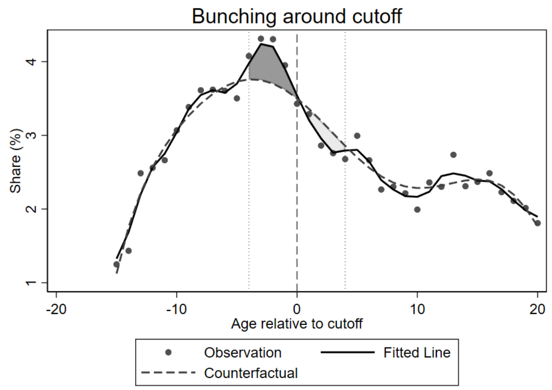
Notes: This figure compares female patients' age-at-first-surgery distributions around the age-31 cutoff across regions with different fertility norms. Fertility norms are measured using the 2010 CFPS question "It is important to have children to carry on the family line?", with answers on a 1–5 scale from "strongly disagree" to "strongly agree". We compute the province-level mean response and classify provinces as having strong or weak fertility norms depending on whether this mean is above or below the national average. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.

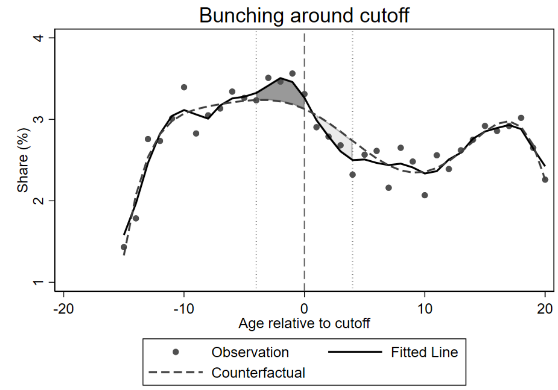


High Intensity, $B=8.03^{***}$ **(2.15)** **Low Intensity,** $B=8.27^{***}(1.46)$

Figure A8: Female age distribution across regions with different Confucian culture

Notes: This figure plots the distribution of women's surgery-age choices relative to the age cutoff in regions with different Confucian culture. The culture intensity is constructed as the density of clan of each prefecture. We classify provinces into strong culture and weak culture groups based on whether this clan density is above or below the national mean. Blue dots show the observed distribution, the solid black line is a smoothed fit, and the red dashed line is the estimated counterfactual distribution in the absence of bunching. The shaded area to the left of the cutoff in each panel represents the excess mass B, which is similar in strong culture regions (left panel) and in weak culture regions (right panel).

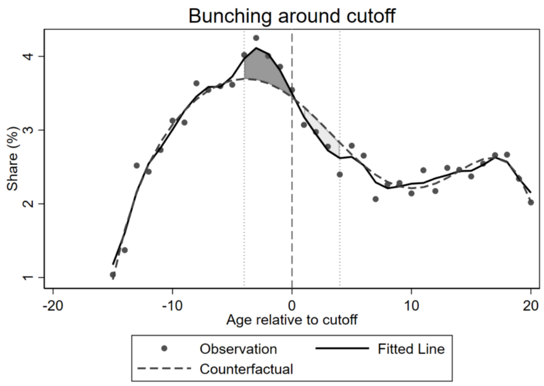




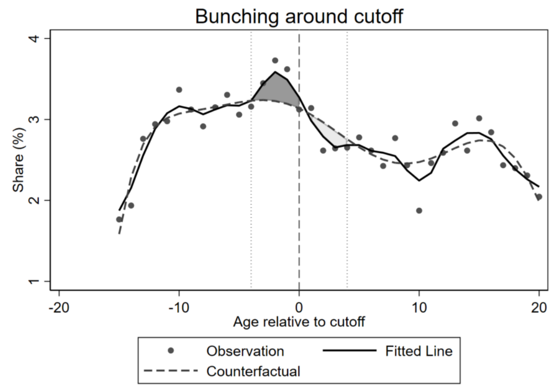
High Ratio, $\hat{B} = 9.49^{***}(1.66)$ **Low Ratio,** $\hat{B} = 6.69^{***}(1.72)$

Figure A9: Female Age Distribution by Local Marriage Market Competition

Notes: This figure compares female patients' age-at-first-surgery distributions around the age-31 cutoff across regions with different marriage pressure. Marriage pressure is constructed from 2018 Baihe.com user data as the ratio of active female users to active male users. Regions are classified as high-pressure or low-pressure depending on whether this ratio is above or below the national mean. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.



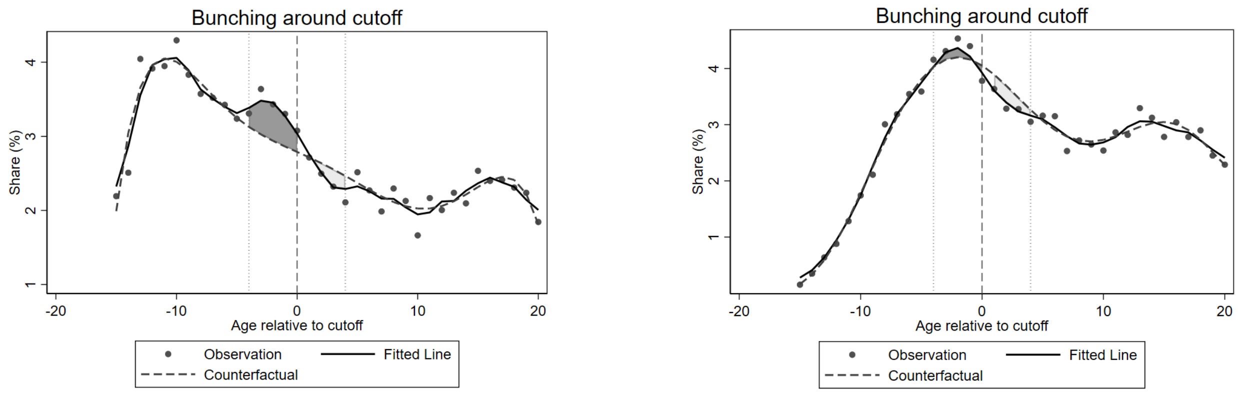
High Intensity, $\hat{B} = 9.26^{***}(1.48)$



Low Intensity, $\hat{B} = 6.56^{***}(2.18)$

Figure A10: Female Age Distribution by Perceived Marriage Pressure

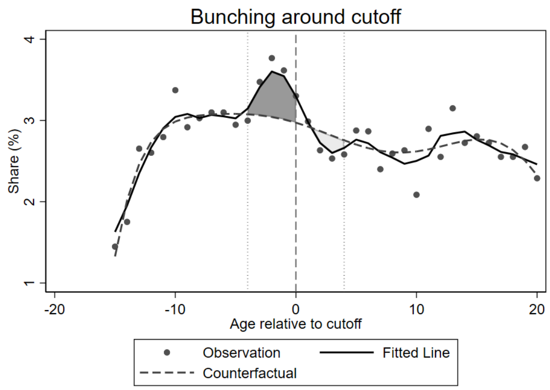
Notes: This figure compares female patients' age-at-first-surgery distributions around the age-31 cutoff across regions with different levels of perceived marriage pressure. Perceived marriage pressure is measured by the Baidu Search Index, constructed as the city-level annual search volume for marriage-pressure (*cuihun*) keywords, averaged over 2017–2019. Cities are classified as high-intensity or low-intensity depending on whether their search frequency is above or below the national mean. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.



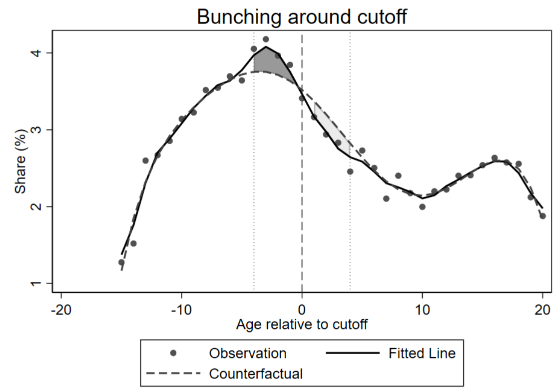
High Dependence, $\hat{B} = 13.66^{***}(2.20)$ **Low Dependence,** $\hat{B} = 2.80^{**}(1.31)$

Figure A11: Female Age Distribution by Economic Dependence

Notes: This figure compares female patients' age-at-first-surgery distributions around the age-31 cutoff across groups with different levels of economic dependence. All cases are cosmetic surgeries. Women are classified as having low or high economic dependence based on occupation information, as described in the text. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.



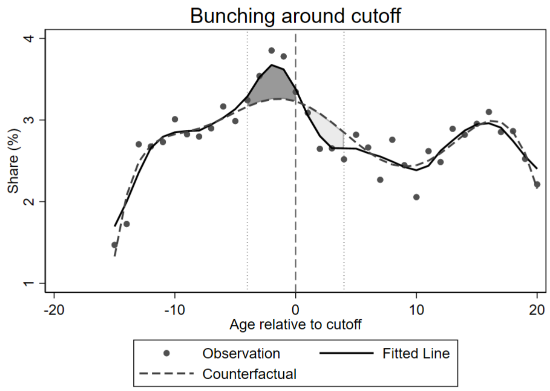
Low Education, $\hat{B} = 13.12^{***}(2.37)$

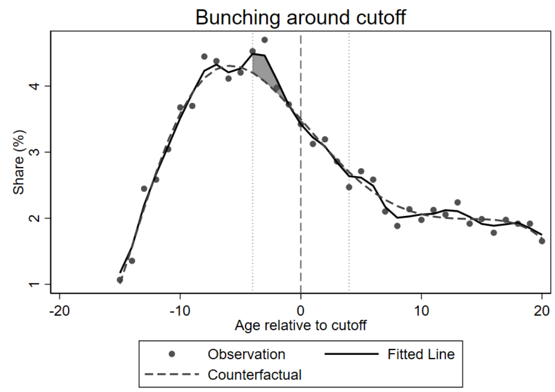


High Education, $\hat{B} = 5.83^{***}(1.12)$

Figure A12: Female Age Distribution by Regional Female Educational Attainment

Notes: This figure compares female patients' age-at-first-surgery distributions around the age-31 cutoff across regions with different levels of female educational attainment. All cases are cosmetic surgeries. Regional female educational attainment is measured using the 2015 Population Census and is defined as the average education level of women aged 25–35 in each province. Provinces are classified as high-education or low-education depending on whether this measure is above or below the national mean. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.





High Gap, $\hat{B} = 10.09^{***}(1.70)$ **Low Gap,** $\hat{B} = 5.00^{***}(1.52)$

Figure A13: Female Age Distribution by Regional Gender Wage Gap

Notes: This figure compares female patients' age-at-first-surgery distributions around the age-31 cutoff across regions with different gender wage gaps. All cases are cosmetic surgeries. The provincial gender wage gap is constructed from the 1% sample of the 2015 China Census combined with industry-level wage data from the National Statistics Data (2017–2019), and is defined as the difference between male and female weighted average wages across industries, using gender-specific employment shares as weights. The gap is larger where men are more concentrated in high-wage sectors. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.

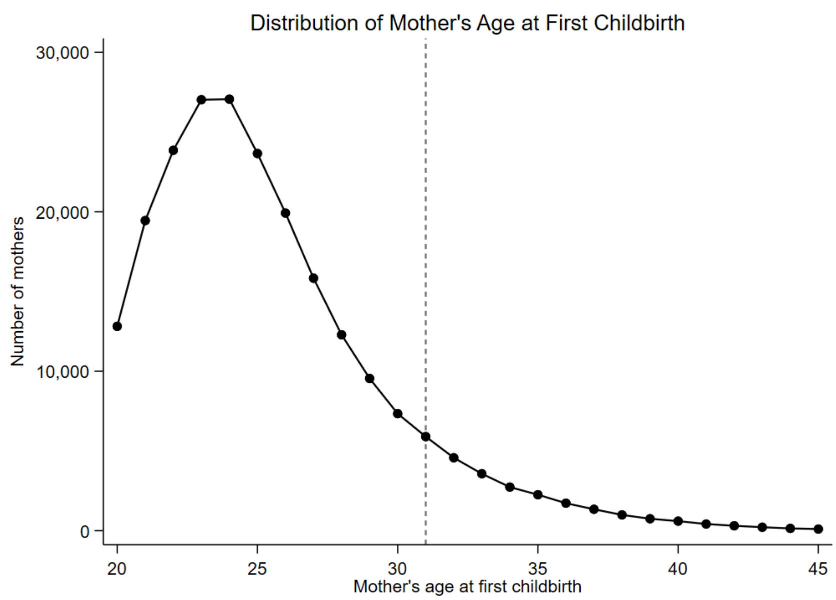


Figure A14: Distribution of Mother's Age at First Childbirth

Notes: This figure plots the distribution of mothers' age at first childbirth using 2015 census data. The horizontal axis reports age at first childbirth, and the vertical axis reports the corresponding frequency of mothers having their first child at each age.

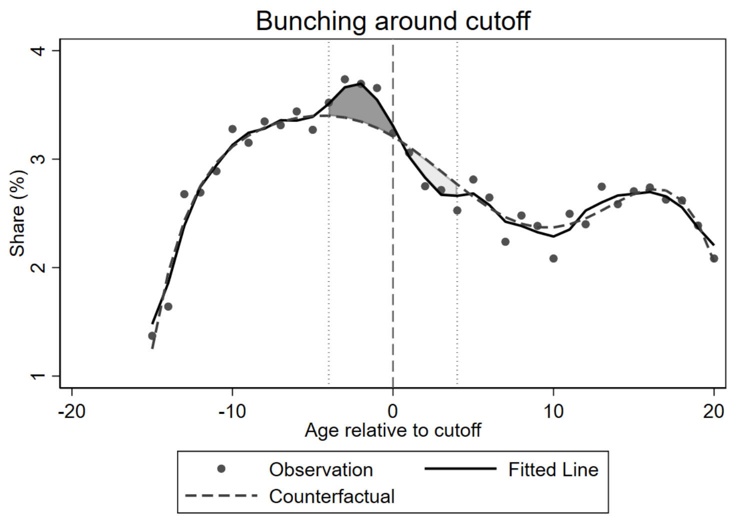


Figure A15: Robustness: Age Distribution after Excluding Prior-Childbirth Cases $\hat{B} = 7.34^{***}(1.42)$

Notes: This figure plots female patients' age-at-first-surgery distribution around the age-31 cutoff, using all surgery types after excluding prior childbirth cases. A case is classified as prior childbirth if the patient name and date of birth match a delivery hospitalization record, the delivery hospitalization occurred before the cosmetic surgery visit, and the delivery and cosmetic surgery hospitals were located in the same province. The horizontal axis reports age relative to the cutoff, and the vertical axis reports the share of surgeries in each one-year age bin. Dots denote the observed distribution, the solid line plots the fitted age profile, and the dashed line plots the counterfactual distribution estimated using Equation 1. The dark shaded area indicates the estimated excess mass immediately to the left of the cutoff, while the light shaded area indicates the corresponding missing mass to the right of the cutoff. The reported $\hat{B}$ is the estimated bunching magnitude, with standard errors in parentheses.

Table A1: Explanation of Randomly Assigned Variables in CGSS Survey

Variable	Assignment Rule
A2	Gender of participant.
X1	Generate the following random variable three times, sequentially for B101–B103. If the respondent is female ($A2 = 2$), draw a random integer n from the interval $[-5, +15]$. If the respondent is male ($A2 = 1$), draw a random integer n from the interval $[-15, +5]$. If $n < 0$, display "$\|n\|$ years younger than you"; if $n > 0$, display "n years older than you"; if $n = 0$, display "the same age as you."
X2	Randomly assign one of the following nine categories: 50% of your income (one-half), 60% of your income, 70% of your income, 80% of your income, 90% of your income, about the same as your income, 150% of your income (1.5 times), 200% of your income (2 times), or 300% of your income (3 times).
X3	Randomly assign one of the following two categories: "parents living in a rural area" or "parents living in an urban area."
X4	Randomly assign one of the following two categories: "does not own a home" or "owns a home."
X5	Randomly assign one of the following three education categories: "high school," "bachelor's degree," or "graduate degree."
X6	Randomly assign one of the following three appearance categories: "somewhat unattractive," "average," or "fairly attractive/fairly handsome." If the respondent is female ($A2 = 2$), the third category is displayed as "fairly handsome"; if the respondent is male ($A2 = 1$), the third category is displayed as "fairly attractive."

Notes: This table describes the randomized attributes used in the marriage-attractiveness experiment in the 2021 wave of the Chinese General Social Survey (CGSS). The variables $X1$–$X6$ correspond to the randomized characteristics of a hypothetical potential partner shown to the respondent in each round. Respondents were asked to evaluate the marital attractiveness of each profile after observing these randomly assigned attributes.

Table A2: Summary Statistics

	Full sample	Male	Female
	(1)	(2)	(3)
Panel A. Surgery Type			
For Beauty	0.472	0.318	0.629
For Functionality	0.436	0.605	0.265
Others	0.092	0.077	0.106
Panel B. Age cohort			
10–19	0.349	0.491	0.205
20–29	0.224	0.192	0.257
30–39	0.162	0.112	0.212
40–49	0.146	0.105	0.187
50–59	0.119	0.100	0.139
Panel C. Body part			
Breast	0.144	0.067	0.223
Ear	0.023	0.030	0.015
Eye	0.261	0.243	0.268
Face	0.088	0.073	0.104
Fat	0.043	0.034	0.053
Genital	0.094	0.184	0.003
Limbs	0.048	0.049	0.047
Nose	0.035	0.034	0.035
Skin	0.173	0.148	0.191
Other Parts	0.115	0.137	0.092
Panel D. Hospitalization period (days)			
< 10	0.782	0.764	0.803
10–30	0.201	0.215	0.186
> 30	0.015	0.020	0.011
Panel E. Hospitalization costs (RMB)			
$< 1k$	0.009	0.007	0.011
$1k - 10k$	0.677	0.680	0.675
$> 10k$	0.314	0.313	0.314
Total observations	102,699	51,726	50,973

Notes: This table summarizes the main characteristics of the plastic surgery sample for the full sample and by gender. Panel A reports the distribution of procedure functionality, distinguishing between surgeries performed primarily for cosmetic reasons ("For Beauty"), for medical or functional indications ("For Functionality"), and a residual "Others" category, which are not included in our following analysis. Panel B presents the age distribution of patients in five ten–year cohorts. Panel C shows the distribution of the main body part targeted by the procedure (e.g., breast, eye, skin). Panel D summarizes the distribution of hospitalization length, grouped into three categories of days spent in hospital for the focal surgery. Panel E summarizes the distribution of hospitalization costs, grouped into three categories of total costs spent in hospital. Entries are shares of observations in each category within the corresponding column.

Table A3: Marriage- and Employment-Related Pressure and Cosmetic Surgery

	Dep. var.: Cosmetic Surgery Count (in log)					
	Marriage-related Index			Employment-related Index		
	Maternal Age (1)	Urged Marriage (2)	Aggregate (3)	Recruitment (4)	Employment (5)	Aggregate (6)
Maternal Age Pressure	1.273**					
	(0.577)					
Urged Marriage Pressure		1.715*				
		(0.906)				
Perceived Marriage Pressure			2.014**			
			(1.018)			
Recruitment Pressure				0.305		
				(0.449)		
Employment Pressure					-0.281	
					(0.952)	
Perceived Employment Pressure						0.370
						(0.592)
City FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Observations	843	843	843	843	843	843

Notes: This table reports the relationship between local pressure indices and cosmetic surgery counts. The dependent variable is the logarithm of the annual number of cosmetic surgery procedures at the city level. Columns (1)–(3) report results for marriage-related pressure measures—maternal age pressure, urged marriage pressure, and the aggregate perceived marriage pressure index. Columns (4)–(6) report results for employment-related pressure measures—recruitment pressure, employment pressure, and the aggregate perceived employment pressure index. The aggregate indices are constructed from the individual search terms by principal component analysis. All specifications include city fixed effects, year fixed effects, and the baseline set of controls. Standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

B Sample Selection

We conduct our sample selection following the updated description of ICD-10 codes of each inpatient record's diagnosis.[14] According to Table B1, we identified 12 categories of plastic surgeries based on the descriptions, and the remaining cases are combined into the category of "Other Parts".

Among the selected plastic-surgery-related cases, we filter out all functional cases and the rest cases are cosmetic. We define a case as functional if its diagnosis code belongs to the genital (i.e. codes N47, N48, Q55, N90, Q52), limb (i.e. codes Q66, Q68, Q72, M95, M20), or metabolic-disorder categories (i.e. code E88). Under other diagnosis codes, cases may include both functional and cosmetic conditions. To identify additional functional plastic-surgery-related cases, we further screened diagnosis notes using a predefined set of keywords:

1. For the face and head region, we included keywords such as cleft lip, cleft palate, ptosis, strabismus, deviated nasal septum, facial atrophy, malocclusion, jaw deviation, cranial defects, and facial palsy.

2. For the chest and torso, we included terms related to funnel chest, pigeon chest, chest wall deformities, thoracic abnormalities, and spinal defects.

3. For the extremities, we included keywords such as inversion, eversion, finger deformities, hand abnormalities, foot abnormalities, clubfoot, upper-limb defects, tibial abnormalities, fibular abnormalities, and muscle deformities.

4. For the genital and breast region, we included terms such as vaginal abnormalities, genital malformations, testicular conditions, scrotal abnormalities, and gynecomastia.

5. We also included a set of general medical terms that may indicate reconstructive or medically necessary treatment, including defect, absence, fixation, complications, sclerosis, ventilation, drug-related injury, metabolic disorders, and pathological changes.

[14] See https://www.icd10data.com/ICD10CM/Codes for detailed descriptions of diagnosis codes

Table B1: ICD-10 Diagnosis Codes for Plastic Surgery

Body part	Codes	Example Keywords
Eye	[H02]	[eye, eyelid, eye bag, double eyelid, eye corner, eyebrow]
Nose	[M95, Q67, Q30]	[nose, nasal bridge, nasal wing, nasal tip, rhinoplasty]
Breast	[N62, Q56, N64, Q83]	[breast, chest, breast augmentation, breast implant]
Face	[K07, Q75, Q87, M89, K13, Q38, Q35, Q36]	[face, facial contour, jaw, maxilla, mandible, cheekbone, chin]
Fat	[E66, E65, E88]	[fat, liposuction, body contouring, waist, abdomen, thigh]
Genital	[N47, N48, Q55, N90, Q52]	[genital, penis, labia, vagina, foreskin]
Skin	[L90, L91, L81, L70, L71, L57, D22, Q82]	[skin, scar, mole, nevus, acne, pigmentation]
Ear	[Q17, H61]	[ear, auricle, earlobe]
Smell	[L75]	[body odor, axillary odor]
Limbs	[Q66, Q68, Q72, M95, M20]	[limb, arm, leg, limb deformity, finger, toe]
Other	[Z41]	[plastic, cosmetic, non-medical, implant]

Notes: This table reports the ICD-10 diagnosis codes and example diagnosis-note keywords used to identify plastic-surgery-related inpatient records and classify them by body part. Cases are assigned to body-part categories based on diagnosis codes and text descriptions in the diagnosis notes. Records that are plastic-surgery-related but cannot be mapped to a specific body-part category are classified as Other. Functional or medically necessary cases are subsequently excluded from the cosmetic-surgery sample using diagnosis codes and keyword-based screening described in the text.

C The "Leftover Women" Age Cutoff

The CGSS 2021 survey presented each participant with 3 profiles of potential marriage partners, describing them across six different dimensions including appearance, housing, age, income, family background, and education. The detailed randomization of each dimension follows the description in Table A1. Participants were asked to score each virtual partner using 7-scale grades.

To explore whether there exists a specific age cutoff that discretely dampens women's marriage prospects, we examine the pure impact of women's age on their perceived attractiveness. The first step in doing so is to control all other available observables other than women's age. Specifically, we employ a TWFE regression:

$$Score_{it} = \alpha + \sum_{s=1}^{5} \beta_s Grade_{it}^{s} + \lambda_i + \tau_t + \epsilon_{it}, \tag{3}$$

where $Score_{it}$ is the score of participant i given at experiment round t for the attractiveness of the given profile. $Grade_{it}^{s}$ is the randomly generated grade for dimension s in participant i's round t. We include the remaining five dimensions of the profile to isolate the effect of age on perceived attractiveness. We control for the individual fixed effect λ_i and round fixed effect τ_t. Standard errors are clustered on the individual level.

To examine the pure effect of age on marriage market prospects, we first extract the residuals from Equation 3. We use the residualized attractiveness score as the adjusted attractiveness measure, capturing the variation in perceived attractiveness that is unexplained by the virtual partner's other five attributes, the respondent's characteristics, and the fixed effects. Next, we calculate the moving average of $\widehat{Score}$ for each specific age of the virtual partners. This smoothing procedure reduces idiosyncratic noise at each exact age and helps reveal the underlying age pattern around the age cutoff more clearly. Then, we run a standard quadratic polynomial regression to quantify the drop at age 31.

In Figure A2, we plot the average $\widehat{Score}$ based on the virtual partners' age, separately for male and female participants. This figure reveals a stark gender asymmetry: the left panel shows that male respondents' ratings of female profiles experience a sharp, discrete drop precisely at age 31 of the female virtual partner. In contrast, the right panel demonstrates that female respondents' ratings of male virtual partners do not suffer a similar penalty at this threshold.

D Adjustment of Counterfactual Distribution

Following Chetty et al. (2011), we apply the method of shifting counterfactual distribution to adjust for the integration constraints. In our context, bunching is interpreted as a reshuffling of the timing of cosmetic surgery within a narrow critical age range around the cutoff. Women who would otherwise undergo surgery shortly after the perceived age threshold may choose to move the timing forward in order to avoid the marriage-market penalty associated with crossing that threshold. Therefore, the relevant mass-balance condition should hold within the bunching window.

When we apply the alternative candidate cutoffs, the estimated counterfactual distribution does not always naturally satisfy this integration constraint. This is particularly the case for candidate cutoffs before age 31, where the raw counterfactual distribution may imply an imbalance between the estimated excess mass and missing mass within the chosen window. To maintain comparability across candidate cutoffs, we therefore adjust the counterfactual distribution following the spirit of Chetty et al. (2011). For cutoffs that do not naturally satisfy the mass-balance condition, we shift the estimated counterfactual distribution upward within the bunching window until the excess mass and missing mass are balanced.

Specifically, let $M = [\hat{a} - h, \hat{a} + h]$ denote the bunching window around a candidate cutoff $\hat{a}$, where h is the half-window width. For each age $a \in M$, let F_a denote the observed distribution and $\widehat{F}_a$ denote the estimated counterfactual distribution. We construct a balanced counterfactual distribution by adding a constant shift κ to the estimated counterfactual distribution within the bunching window:

$$\widehat{F}_a^{bal} = \widehat{F}_a + \kappa, \qquad a \in M.$$

The shift κ is chosen to eliminate the average discrepancy between the observed and counterfactual distributions within the bunching window:

$$\kappa = \frac{\sum_{a \in M} \left(F_a - \widehat{F}_a \right)}{2h + 1}.$$

This adjustment ensures that the balanced counterfactual distribution satisfies the within-window integration constraint:

$$\sum_{a \in M} \left(F_a - \widehat{F}_a^{bal} \right) = 0.$$

E Calculating Gender Wage Gap

Using the 1% population sample of the 2015 China Census, we compute the gender wage gap for each province. The 1% census sample is a representative dataset that records detailed demographic and labor market information. We calculate the gender wage gap following:

$$Gap_p = \sum_i (s_{ip}^m \times w_{ip}^m) - \sum_i (s_{ip}^f \times w_{ip}^f), \tag{4}$$

where $s_{ip}^{m,f}$ is the male or female average income for industry i and province p. The industry×province wage is provided by National Statistics Data. The $w_{ip}^{m,f}$ is the share of employed male or female of industry i in province p among all male or female employees in province p. The gender wage gap is essentially the gap of weighted average wage of male and female.

The intuition of this method is that the gender wage gap typically stems from the selection process. Specifically, when there is a higher proportion of male workers in the high-waged industries, the weighted average wage would be higher for male. Moreover, we calculate the wage gap under an assumption that besides selection, there is no difference and discrimination in the compensation schemes for male and female in every industry.

F Model: Setup, Derivations, and Proofs

F.1 Setup

Consider a continuum of single women who differ in their marriage-market appeal outside physical attractiveness. We summarize these non-beauty attributes by a type $\theta \in \Theta \subset \mathbb{R}$, distributed according to F with a smooth density. The type can be read as a reduced-form index of education, earnings potential, occupational status, family background, and other characteristics rewarded in the marriage market.

At time $t = 0$, a woman chooses a surgery plan: either select a single age $\tau \in [t_0, T]$ at which to undergo cosmetic surgery, or forgo surgery altogether. Surgery is binary, irreversible, and costly. If she remains single until τ, she pays a fixed cost $c > 0$; if she matches before then, the procedure is never undertaken and the cost is not incurred. This timing structure captures the key option value in the problem: postponing surgery preserves the possibility that the investment becomes unnecessary because a match arrives naturally. We impose commitment to keep the model transparent; the same mechanism would arise in a sequential timing problem.

The age-specific matching hazard takes the form

$$\lambda(t; \tau, \theta) = \lambda_0 + \mathbf{1}[t \geq \tau] \cdot \Delta(\tau; \theta),$$

Before the procedure, the woman matches at her baseline hazard $\lambda_0 > 0$, a feature of the marriage market she faces rather than of her type; Corollary 2 studies how market thickness shifts λ_0. Once surgery is completed, the matching hazard increases by

$$\Delta(\tau; \theta) = m\, \phi(\theta)\, g(\tau)\, \rho(\tau; \theta), \qquad \rho(\tau; \theta) = \begin{cases} 1 & \tau \leq \bar{t}, \\ \kappa(\theta) & \tau > \bar{t}. \end{cases} \tag{5}$$

The boost is the product of a procedure-level salience, a type-level return, a life-cycle profile, and the norm multiplier, each corresponding to a distinct economic force.

The scalar $m > 0$ is the *beauty salience of the procedure*: how strongly it raises a woman's marriage-market attractiveness. It is a property of the procedure, common across women, and indexes procedures in the cross-section—facial and breast surgeries carry high m, more diffuse or functional procedures low m. Within the baseline model (a single procedure) m is a fixed scalar; Corollary 1 varies it across procedures.

The function $\phi(\theta) > 0$ is the *type-θ woman's marriage-market return to a unit of surgery-induced beauty.* It nets together the scale of her partner pool and her responsiveness to appearance; because these enter the hazard only through their product, they are not separately identified, and we work with the composite ϕ directly. Because beauty and non-beauty appeal are substitutes in producing marital prospects, this return is smaller for women already strong on other dimensions, so $\phi'(\theta) < 0$.

The function $g(\tau)$ captures how the marriage-market salience of appearance enhancement evolves over the life cycle. We assume only that g is strictly increasing and log-concave: the return to surgery rises with age, but at a decelerating *proportional* rate. We do not impose any interior peak. The eventual unattractiveness of further delay does not come from a decline in g; it comes from discounting and from the competing risk of matching before the procedure, both of which are embedded in the survival-weighted objective derived below.

The norm multiplier $\rho(\tau;\theta)$ introduces the age-based stigma. Surgery undertaken at or before $\bar{t}$ receives the full return ($\rho = 1$). Surgery deferred beyond $\bar{t}$ receives only a fraction $\kappa(\theta) < 1$ of that return. Economically, crossing the threshold reduces the marriage-market payoff to appearance enhancement: once a woman is viewed through the "leftover" lens, the same physical improvement generates a smaller gain in perceived desirability. The penalty is allowed to be weaker for women with stronger non-beauty appeal, whose other attributes can partially offset the stigma; hence $\kappa'(\theta) \geq 0$. The step-function form is a stylized representation of the discrete decline in perceived female attractiveness documented in Section 4.1.

A match delivers lifetime value V^M, and future payoffs are discounted at rate $r > 0$. We set $T = \infty$ for tractability; allowing a finite horizon would not change the mechanism. The woman chooses τ—or chooses never to undergo surgery—to maximize expected discounted utility.

Assumptions.

(A1) Regularity. $\phi, \kappa, g \in C^2$; $m > 0$ and $c > 0$ are fixed scalars; g is strictly increasing ($g'(\tau) > 0$) and log-concave ($(\log g)'' < 0$, equivalently g'/g strictly decreasing) on $[t_0, T]$.

(A2) Substitution. $\phi'(\theta) < 0$.

(A3) Norm asymmetry. $\kappa(\theta) < 1$ for all θ, with $\kappa'(\theta) \geq 0$.

F.2 Payoff

We next express the woman's payoff under a given surgery plan. If she chooses age τ, the probability of still being single at time t is

$$S(t) = \begin{cases} e^{-\lambda_0 t}, & t \le \tau, \\ e^{-\lambda_0 \tau}\, e^{-(\lambda_0+\Delta)(t-\tau)}, & t > \tau, \end{cases}$$

where $\Delta \equiv \Delta(\tau;\theta)$. The first line is survival before surgery; the second line multiplies survival up to τ by survival under the elevated post-surgery hazard.

Direct integration of the discounted match flow over $[0,\infty)$ yields

$$\begin{aligned} \int_0^\infty e^{-rt}\lambda(t)V^M S(t)\,dt &= \lambda_0 V^M \int_0^\tau e^{-(r+\lambda_0)t}dt + (\lambda_0+\Delta)V^M e^{-\lambda_0\tau}\int_\tau^\infty e^{-rt}e^{-(\lambda_0+\Delta)(t-\tau)}dt \\ &= \frac{\lambda_0 V^M}{r+\lambda_0}\big(1-e^{-(r+\lambda_0)\tau}\big) + \frac{(\lambda_0+\Delta)V^M}{r+\lambda_0+\Delta}\,e^{-(r+\lambda_0)\tau}. \end{aligned}$$

The first term is the value of matching before surgery under the baseline hazard. The second term is the continuation value for women who reach τ unmatched and then search with the higher hazard $\lambda_0+\Delta$.

The expected discounted cost is $e^{-r\tau}S(\tau)c = e^{-(r+\lambda_0)\tau}c$, since the cost is paid only by women who remain single until τ. Combining benefits and cost, and defining the no-surgery baseline value

$$V_1 := \frac{\lambda_0 V^M}{r+\lambda_0},$$

algebraic simplification yields

$$U(\tau;\theta) = V_1 + e^{-(r+\lambda_0)\tau}\,\Phi(\tau;\theta), \tag{6}$$

where the *surgery surplus* is

$$\Phi(\tau;\theta) = \frac{rV^M\,\Delta(\tau;\theta)}{(r+\lambda_0)\big(r+\lambda_0+\Delta(\tau;\theta)\big)} - c. \tag{7}$$

Verification. The incremental match value from the elevated hazard is

$$\frac{(\lambda_0+\Delta)V^M}{r+\lambda_0+\Delta} - \frac{\lambda_0 V^M}{r+\lambda_0} = V^M\cdot\frac{(\lambda_0+\Delta)(r+\lambda_0)-\lambda_0(r+\lambda_0+\Delta)}{(r+\lambda_0+\Delta)(r+\lambda_0)} = \frac{rV^M\Delta}{(r+\lambda_0)(r+\lambda_0+\Delta)},$$

which gives (7).

Equation (6) separates the problem into a baseline value and an investment value. The baseline V_1 is the payoff from never undergoing surgery and does not vary with the type or the plan. The second term is the incremental value of the surgery plan: the probability of arriving at τ unmatched, discounted by $e^{-(r+\lambda_0)\tau}$, multiplied by the net surplus $\Phi(\tau;\theta)$ available at that age. Since V_1 does not vary with τ, the timing problem is to maximize $e^{-(r+\lambda_0)\tau}\Phi(\tau;\theta)$. Surgery is chosen only if the best attainable surplus is positive.

The surplus Φ has the usual investment interpretation. The gross benefit rises with the hazard increment Δ and with the value of marriage V^M, while the denominator adjusts for discounting and for the competing risks of matching before and after surgery. The fixed cost c shifts the surplus downward. The gain is increasing but concave in Δ, so the value of additional appearance-induced matching improvements is naturally bounded.

F.3 First-order Condition for Optimal Timing

We now characterize the optimal surgery age. Let $W(\tau;\theta) := V_1 + e^{-(r+\lambda_0)\tau}\Phi(\tau;\theta)$. Differentiating with respect to τ gives

$$\frac{\partial W}{\partial \tau} = e^{-(r+\lambda_0)\tau}\left[\frac{\partial \Phi}{\partial \tau} - (r+\lambda_0)\Phi\right].$$

On a segment where ρ is constant (i.e., $\tau \neq \bar{t}$), the chain rule gives

$$\frac{\partial \Phi}{\partial \tau} = \frac{\partial \Phi}{\partial \Delta}\cdot\frac{\partial \Delta}{\partial \tau} = \frac{rV^M}{(r+\lambda_0+\Delta)^2}\cdot m\,\phi(\theta)\,\rho(\tau;\theta)\,g'(\tau),$$

where $\partial\Phi/\partial\Delta = rV^M/(r+\lambda_0+\Delta)^2$ and $\partial\Delta/\partial\tau = m\phi\rho g'(\tau)$. The first-order condition is therefore

$$\underbrace{(r+\lambda_0)\Phi(\tau;\theta)}_{\text{flow cost of delay}} = \underbrace{\frac{rV^M m\,\phi(\theta)\,g'(\tau)\rho(\tau;\theta)}{[r+\lambda_0+\Delta(\tau;\theta)]^2}}_{\text{flow gain from delay}}. \tag{FOC}$$

The timing choice balances the cost of waiting against the benefit of waiting. The left-hand side is the flow cost of delay: future surplus is discounted at rate r, and each additional instant of waiting also creates the possibility that the woman matches before investing, in which case the surgery surplus is never realized. The right-hand side is the flow gain from delay. It is proportional to $g'(\tau)$, so waiting is attractive only on ages where the marriage-market salience of surgery is still rising.

Compact form. Substituting $m\phi\rho = \Delta/g(\tau)$ (from (5)) and using the identity

$$\frac{rV^M\Delta}{(r+\lambda_0+\Delta)^2} = \frac{rV^M\Delta}{(r+\lambda_0)(r+\lambda_0+\Delta)}\cdot\frac{r+\lambda_0}{r+\lambda_0+\Delta} = (\Phi+c)\cdot\frac{r+\lambda_0}{r+\lambda_0+\Delta},$$

the (FOC) can be rewritten as

$$\Phi^*(\tau;\theta)\big[r+\lambda_0+\Delta^*(\tau;\theta)\big] = \big[\Phi^*(\tau;\theta)+c\big]\cdot\frac{g'(\tau)}{g(\tau)}. \tag{8}$$

This representation makes the timing force especially clear. The left-hand side is net surplus scaled by the effective exit rate, while the right-hand side is the gross benefit multiplied by the logarithmic growth rate of salience, $g'(\tau)/g(\tau)$. Since surgery is undertaken only when $\Phi^* > 0$ and g is strictly increasing, the right-hand side is positive, so an interior optimum is consistent with $g'(\tau) > 0$ everywhere. Optimal timing is pinned down by the balance between the still-rising salience and the effective exit rate $r+\lambda_0$, rather than by any turning point of g. Because g'/g is strictly decreasing under log-concavity, this balance is met at a unique age, as the next lemma shows.

Lemma 1 (Single-peakedness of the timing objective)**.** *Fix θ and a segment on which the norm multiplier ρ is constant. Under (A1), $W(\cdot;\theta)$ is strictly quasi-concave on that segment, so its maximizer is unique.*

Proof. On the segment write $\Delta(\tau) = \beta\, g(\tau)$ with $\beta \equiv m\,\phi(\theta)\,\rho > 0$, and set $a \equiv r+\lambda_0 > 0$ and $\gamma(\tau) \equiv g'(\tau)/g(\tau)$. From (7), $\Phi = G(\Delta) - c$ with

$$G(\Delta) = \frac{rV^M}{a}\cdot\frac{\Delta}{a+\Delta}, \qquad G'(\Delta) = \frac{rV^M}{(a+\Delta)^2} > 0.$$

Since $\beta\, g' = \Delta\,\gamma$, the marginal payoff satisfies

$$e^{a\tau}\frac{\partial W}{\partial \tau} = \Phi'(\tau) - a\Phi(\tau) = a\Big[\,G(\Delta)\,\underbrace{\Big(\tfrac{\gamma}{a+\Delta}-1\Big)}_{=:B(\tau)} + c\,\Big].$$

By log-concavity γ is strictly decreasing, and Δ is strictly increasing, so B is strictly decreasing; meanwhile $G(\Delta) > 0$ is strictly increasing. At any critical point the bracket vanishes, i.e. $G(\Delta)\,B = -c < 0$, which forces $B < 0$. Wherever $B < 0$ we have $\frac{d}{d\tau}[G(\Delta)B] < 0$ (the first factor has $G'_\tau > 0$, $B < 0$; the second has $G > 0$, $B'_\tau < 0$), so $\partial W/\partial\tau$ is strictly decreasing through every critical point. Hence W admits a single interior critical point, which is a maximum (or attains its maximum at a corner), and is strictly quasi-concave. □

Segment maxima. The threshold $\bar{t}$ divides the timing problem into two smooth pieces. On $[t_0,\bar{t}]$, the beauty return is undiminished; on $(\bar{t},T]$, the same procedure receives only the scaled return $\kappa < 1$. We therefore define the best timing on each side of the threshold as

$$\tau_L^*(\theta) := \arg\max_{\tau \in [t_0, \bar{t}]} W(\tau;\theta), \qquad \tau_R^*(\theta) := \arg\max_{\tau \in [\bar{t}, T]} W(\tau;\theta).$$

By Lemma 1, W is strictly quasi-concave on each open segment, so the segment optima are unique. The key case is when the left-segment optimum occurs at the corner, $\tau_L^*(\theta) = \bar{t}$. Such women would like to move further right if the post-threshold penalty were absent, but crossing the threshold reduces the payoff to doing so. These are the types for whom bunching can arise.

Intuition. The compact FOC also previews the comparative statics. A higher baseline matching rate λ_0 makes delay more costly and pulls optimal timing earlier. A higher beauty return $m\phi$ raises the hurdle $\Phi(r + \lambda_0 + \Delta)/(\Phi + c)$, which on the decreasing schedule g'/g is met only at earlier ages. The bunching results below aggregate these individual timing responses across types.

F.4 Bunching at the Norm Threshold

Lemma 2. *Under (A1)–(A3), $W(\bar{t};\theta) > W(\bar{t}^+;\theta)$.*

Proof. By (5), $\Delta(\bar{t};\theta;\rho = 1) = m\phi g(\bar{t}) > \kappa m\phi g(\bar{t}) = \Delta(\bar{t};\theta;\rho = \kappa)$ since $\kappa < 1$. The map $\Delta \mapsto \Phi$ in (7) is strictly increasing in Δ on $\Delta \geq 0$, with derivative

$$\frac{\partial \Phi}{\partial \Delta} = \frac{rV^M}{(r + \lambda_0 + \Delta)^2} > 0.$$

Hence $\Phi(\bar{t};\theta;\rho = 1) > \Phi(\bar{t};\theta;\rho = \kappa)$. Note that the cost c enters Φ additively on both sides and therefore cancels in the difference; the gap originates entirely from the gross-benefit term. Therefore

$$W(\bar{t};\theta) - W(\bar{t}^+;\theta) = e^{-(r+\lambda_0)\bar{t}}\big[\Phi(\bar{t};\theta;\rho = 1) - \Phi(\bar{t};\theta;\rho = \kappa)\big] > 0. \qquad \square$$

Economic interpretation of the jump. Lemma 2 is the source of the deadline effect. The lifetime payoff falls discontinuously when surgery is postponed from just before to just after $\bar{t}$, because the same physical investment now receives only the scaled marriage-market return. The expected size of this loss is $e^{-(r+\lambda_0)\bar{t}}\big[\Phi_L(\bar{t}) - \Phi_R(\bar{t})\big]$. The survival term discounts the loss by the probability of remaining single until the threshold, while the surplus gap captures the reduction in the beauty return. When $\kappa(\theta)$ is low, the gap is large and the threshold exerts a stronger pull; when $\kappa(\theta)$ approaches one, the discontinuity disappears.

The bunching boundary τ^{**}. The downward jump creates an initial region of post-threshold ages that are dominated by $\bar{t}$. Define

$$\tau^{**}(\theta) := \inf\{\tau > \bar{t} : W(\tau; \theta) > W(\bar{t}; \theta)\},$$

the first age beyond the threshold at which delaying strictly dominates locating at $\bar{t}$, with the convention $\tau^{**}(\theta) = T$ when no such age exists. By Lemma 2, $W(\bar{t}^{+}) < W(\bar{t})$, so $\tau^{**}(\theta) > \bar{t}$. By strict quasi-concavity on the right segment (Lemma 1), W rises from $W(\bar{t}^{+})$ to its right-segment peak τ_R^* and then declines toward 0; hence either the peak clears $W(\bar{t})$, in which case $\tau^{**} \in (\bar{t}, \tau_R^*)$ is the unique first upcrossing, or it does not, in which case $W(\tau) \leq W(\bar{t})$ throughout $(\bar{t}, T]$ and $\tau^{**} = T$. In both cases, type θ bunches at $\bar{t}$ if and only if $\tau_R^*(\theta) \leq \tau^{**}(\theta)$, equivalently if and only if the right-segment peak fails to clear the threshold payoff. The interval $[\bar{t}, \tau^{**}]$ is the catchment area of the deadline, and the comparative statics below operate by expanding or contracting it.

Proposition 1 (Bunching at the norm threshold)**.** *Under the assumptions stated in Appendix F.1, the distribution of enhancement timing exhibits a mass point at the norm threshold. Bunching is one-sided: only women whose unconstrained optimal timing falls in a right-neighborhood of the threshold relocate to it, whereas women whose unconstrained optimum lies below the threshold do not. Away from the threshold, the timing distribution is smooth.*

Proof. The woman's problem reduces to comparing the two segment maxima:

$$\tau^o(\theta) = \arg\max\{W(\tau_L^*(\theta); \theta),\ W(\tau_R^*(\theta); \theta)\}.$$

We consider the two cases corresponding to whether the left-segment optimum is interior or attained at the corner.

Case 1: $\tau_L^*(\theta) < \bar{t}$. The left-segment optimum is interior, so by strict quasi-concavity of W on $[t_0, \bar{t}]$ (Lemma 1), $W(\tau_L^*) > W(\bar{t})$. The woman picks τ_L^* if $W(\tau_L^*) \geq W(\tau_R^*)$, and τ_R^* otherwise. In neither sub-case does she choose $\bar{t}$.

Case 2: $\tau_L^*(\theta) = \bar{t}$. The left-segment optimum is at the corner, meaning that throughout $[t_0, \bar{t}]$, $\partial W/\partial \tau \geq 0$ by (FOC). Hence $W(\bar{t}) \geq W(\tau)$ for all $\tau \in [t_0, \bar{t}]$. The choice reduces to a binary comparison:

$$\tau^o(\theta) = \begin{cases} \bar{t}, & W(\bar{t}; \theta) \geq W(\tau_R^*(\theta); \theta), \\ \tau_R^*(\theta), & \text{otherwise.} \end{cases}$$

By the construction of $\tau^{**}(\theta)$ above, $W(\tau_R^*) \leq W(\bar{t}) \iff \tau_R^* \leq \tau^{**}$. Hence type θ bunches at $\bar{t}$ if and only if $\tau_R^*(\theta) \in [\bar{t}, \tau^{**}(\theta)]$.

Mass point and continuity. By the implicit function theorem applied to (FOC), τ_L^* and τ_R^* are continuous in θ. The set $\{\theta : \tau^o(\theta) = \bar{t}\}$ is the union of Case-2 types satisfying $W(\bar{t}) \geq W(\tau_R^*)$, which forms a union of intervals of positive F-measure whenever bunching arises (i.e., whenever the right-segment optimum of any Case-2 type falls inside its catchment $[\bar{t}, \tau^{**}]$). The equilibrium density of τ^o is the pushforward of F through τ^o: a smooth density (from interior Case-1 and large-τ_R^* Case-2 choices) plus a Dirac mass at $\bar{t}$ (from bunching Case-2 types).

Asymmetry. Case-1 types ($\tau_L^* < \bar{t}$) strictly prefer their interior left optimum to $\bar{t}$. Hence no type with unconstrained optimum strictly below $\bar{t}$ relocates to $\bar{t}$: bunching is one-sided. In particular, there is no compensating dip in the density just below the bunching point, unlike in policy-induced bunching settings where the dip arises from agents moving *up* into the threshold from below. □

Intuition behind the asymmetry. The bunching is one-sided because the norm changes the level of payoffs at the threshold rather than merely changing a slope. Women whose preferred timing is already below $\bar{t}$ have no incentive to delay toward the threshold: doing so moves them away from their optimum without avoiding any additional penalty. By contrast, women whose preferred timing lies just beyond $\bar{t}$ face a discrete loss from crossing it. For some of them, moving surgery back to the threshold sacrifices only a small amount of age salience while preserving the beauty return. This level discontinuity generates a mass point at the threshold without requiring a compensating dip below it.

F.5 An Example

The propositions above are qualitative. To fix ideas and visualize the mechanism end to end, this subsection specializes the environment to a concrete parameterization satisfying (A1)–(A3)—for instance a saturating salience profile $g(\tau) = 1 - e^{-\eta\tau}$, which is strictly increasing and log-concave because $g'/g = \eta e^{-\eta\tau}/(1 - e^{-\eta\tau})$ is strictly decreasing, together with a smooth type density F. None of the results of the previous subsections depends on this particular choice; it serves only to generate the figures. One feature of the benchmark schedule is worth noting: the most responsive types (low θ) are sent to the corner t_0, which places a mass point at t_0. Because t_0 lies far below $\bar{t}$, the counterfactual density is smooth in a neighborhood of $\bar{t}$, which is all that the bunching comparison requires.

Figures F1–F4 summarize the mechanism graphically. Figure F1 starts from the no-norm benchmark, in which individual surgery timing $\tau^*(\theta)$ varies smoothly with type. Aggregating over $F(\theta)$ produces the smooth counterfactual density $f_0(t)$ in Figure F2. Figure F3 then introduces the norm: the payoff discontinuity splits the timing problem into two segments and makes the threshold attractive for women whose preferred timing lies just to its right. Finally, Figure F4 shows the implied population density, with excess mass at the threshold and missing mass above it relative to the no-norm counterfactual.

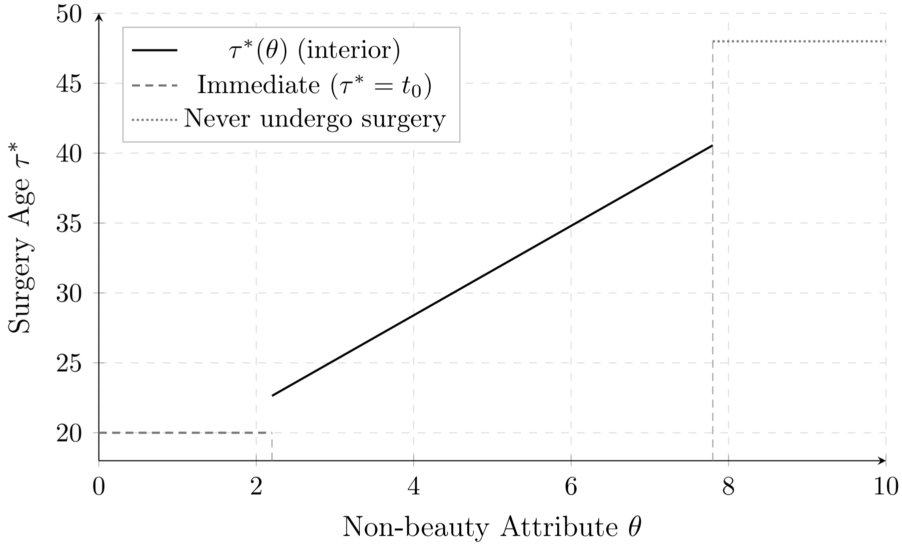


Figure F1: Individual-level Benchmark Schedule $\tau^*(\theta)$

Notes: This figure shows the individual-level optimal surgery age $\tau^*(\theta)$ in the no-norm benchmark ($\rho \equiv 1$). Three regimes emerge under (A1)–(A2): for $\theta < \underline{\theta}$, the beauty return $m\phi(\theta)$ is large and the woman undergoes surgery immediately at the earliest feasible age t_0; for $\theta \in [\underline{\theta}, \bar{\theta}]$, the (FOC) admits an interior solution that is strictly increasing in θ; for $\theta > \bar{\theta}$, the beauty return is too small and no surgery occurs. The schedule is continuous and weakly monotone on the interior region, so aggregating over the type density yields the smooth counterfactual surgery-age density shown in Figure F2; the corner regime places a mass point at t_0, far below $\bar{t}$.

F.6 Comparative Statics

We now use the model to organize the heterogeneity patterns in the data. Let $\mathcal{B}(\rho) := \{\theta : \tau^o(\theta) = \bar{t}\}$ be the set of types who bunch in environment ρ, and let $B(\rho) := \int \mathbf{1}[\theta \in \mathcal{B}(\rho)]\, dF(\theta; \rho)$ denote its mass. The comparative statics work through two objects: the size of the payoff jump at $\bar{t}$ and the location of the right-segment optimum τ_R^* relative to the catchment boundary τ^{**}.

Corollary 1 (Beauty-premium salience)**.** *Bunching is more pronounced for procedures with higher marriage-market salience of beauty.*

Proof of Corollary 1 (beauty-premium salience). Higher procedure salience raises the marriage-market return to surgery. The relevant question is whether it also enlarges the discontinuity at $\bar{t}$, because that discontinuity is what draws women to the threshold.

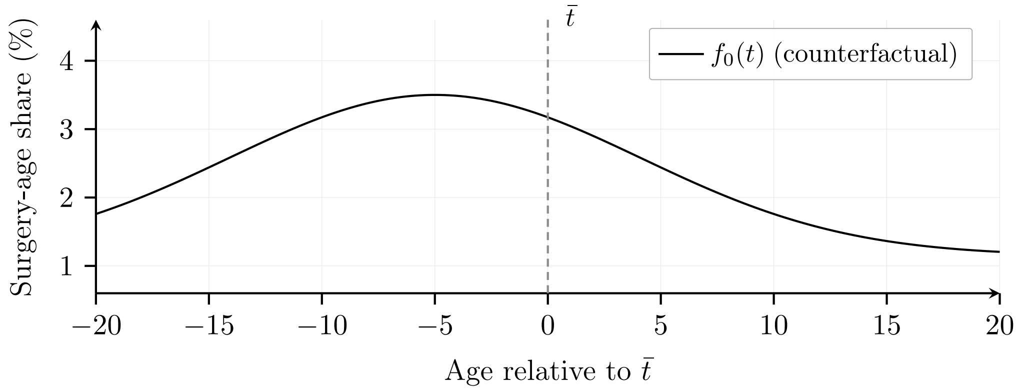


Figure F2: Population-level Counterfactual Surgery-age Density $f_0(t)$

Notes: This figure plots the counterfactual population-level surgery-age density $f_0(t)$, obtained by aggregating the individual schedule $\tau^*(\theta)$ of Figure F1 over the type distribution $F(\theta)$ in the no-norm environment ($\rho \equiv 1$). Because the interior schedule is monotone and F admits a smooth density, f_0 is smooth in a neighborhood of the threshold age $\bar{t}$ (vertical dashed line). This local smoothness of f_0 at $\bar{t}$ is the baseline against which norm-induced bunching is identified in Figure F4.

Define the gross benefit $\mathcal{G}(\Delta) := rV^M\Delta/[(r+\lambda_0)(r+\lambda_0+\Delta)]$, so that $\Phi = \mathcal{G}(\Delta) - c$. Evaluated at $\bar{t}$, the boost equals $\Delta_L = m\phi g(\bar{t})$ on the left and $\Delta_R = \kappa\Delta_L$ on the right. The jump in Φ at $\bar{t}$ is

$$J(\theta; m) := \Phi_L(\bar{t}) - \Phi_R(\bar{t}) = \mathcal{G}(\Delta_L) - \mathcal{G}(\Delta_R),$$

with c cancelling. Differentiating in the procedure salience m,

$$\frac{\partial J}{\partial m} = \mathcal{G}'(\Delta_L) \cdot \frac{\partial \Delta_L}{\partial m} - \mathcal{G}'(\Delta_R) \cdot \frac{\partial \Delta_R}{\partial m} = \phi g(\bar{t})\big[\mathcal{G}'(\Delta_L) - \kappa\mathcal{G}'(\Delta_R)\big],$$

where $\mathcal{G}'(\Delta) = rV^M/(r+\lambda_0+\Delta)^2 > 0$. Since $\Delta_L > \Delta_R$, we have $\mathcal{G}'(\Delta_L) < \mathcal{G}'(\Delta_R)$; whether $\partial J/\partial m > 0$ thus depends on whether $\mathcal{G}'(\Delta_L) > \kappa\mathcal{G}'(\Delta_R)$. Direct algebra (substituting $\mathcal{G}'$ and writing $\Delta_R = \kappa\Delta_L$) shows that this inequality holds whenever

$$(r+\lambda_0+\kappa\Delta_L)^2 > \kappa(r+\lambda_0+\Delta_L)^2,$$

which requires that the effective baseline exit rate, $r+\lambda_0$, not be dominated by the surgery boost. Under this regularity condition, increasing m widens the payoff jump, moves τ^{**} to the right, and expands the bunching set $\mathcal{B}$. Empirically, this comparative static corresponds to comparing procedures with high marriage-market salience, such as facial and breast surgeries (high m), to other cosmetic procedures with more diffuse marital relevance (low m). □

Intuition for Corollary 1. When a procedure raises beauty more strongly, the return to surgery is higher, and the norm removes a larger amount of value once the threshold

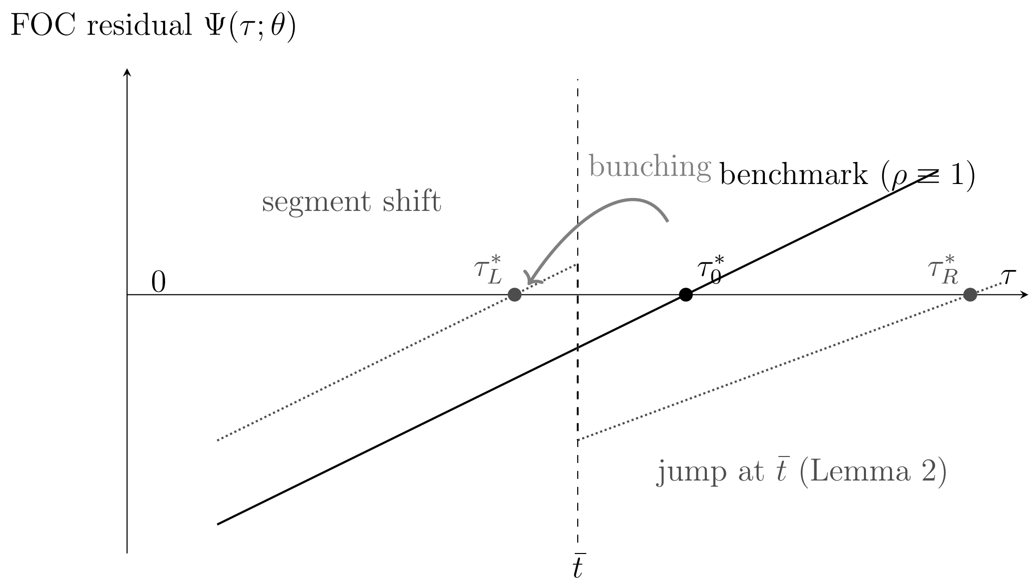


Figure F3: Individual-level FOC Mechanism under the Marriage-age Norm

Notes: This figure illustrates the FOC mechanism schematically. The solid black line shows the FOC residual $\Psi(\tau;\theta)$ in the no-norm benchmark ($\rho \equiv 1$), defined so that $\Psi(\tau^*;\theta) = 0$ at the unconstrained optimum τ_0^* past the threshold $\bar{t}$. Activating the norm ($\rho = \kappa < 1$ past $\bar{t}$) splits the residual into two segments with optima τ_L^* on $[t_0, \bar{t}]$ and τ_R^* on $(\bar{t}, T]$, separated by a downward jump at $\bar{t}$ established in Lemma 2. For Case-2 types whose right-segment optimum τ_R^* lies in the bunching region $[\bar{t}, \tau^{**}]$ of Proposition 1, the discrete penalty exceeds the gain from delay and the woman strictly prefers relocating to the threshold, as indicated by the curved arrow.

is crossed. The threshold therefore becomes a stronger deadline: more women who would otherwise wait slightly longer choose to move the procedure back to $\bar{t}$.

Corollary 2 (Market thickness)**.** *Bunching is more pronounced where the female-side marriage market is thinner.*

Proof of Corollary 2 (market thickness). Market thickness operates through the baseline hazard λ_0 and affects bunching through both surplus and timing.

(a) Surplus and jump shrinkage. The marriage surplus $V^M - V_1 = rV^M/(r+\lambda_0)$ is strictly decreasing in λ_0, and inspection of (7) shows $\Phi(\tau;\theta) \downarrow -c$ uniformly in τ as $\lambda_0 \to \infty$: as matches arrive arbitrarily quickly, the value of any surgery vanishes and at the limit surgery is abandoned, so $B \to 0$. More precisely, the payoff jump is

$$W(\bar{t}) - W(\bar{t}^+) = e^{-(r+\lambda_0)\bar{t}}\big[\mathcal{G}(\Delta_L) - \mathcal{G}(\Delta_R)\big],$$

and *both* factors strictly decline in λ_0: the survival/discount factor $e^{-(r+\lambda_0)\bar{t}}$ trivially, and the gross-benefit gap because $\partial\mathcal{G}/\partial\lambda_0 < 0$ with the larger boost Δ_L the more sensitive, so the gap $\mathcal{G}(\Delta_L) - \mathcal{G}(\Delta_R)$ narrows. The absolute incentive to bunch therefore shrinks as the

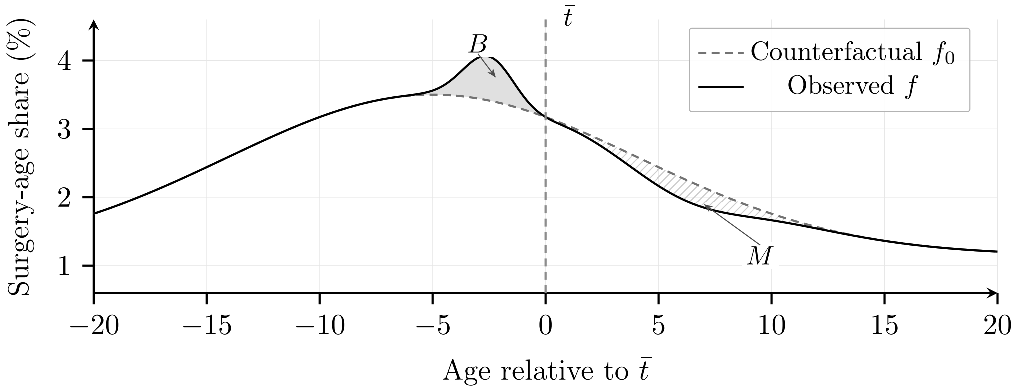


Figure F4: Population-level Bunching under the Marriage-age Norm

Notes: This figure compares the population-level surgery-age density with and without the marriage-age norm. The dashed line reproduces the counterfactual f_0 from Figure F2; the solid line shows the equilibrium density f under $\rho < 1$ past $\bar{t}$. By Proposition 1, types whose right-segment optimum lies in $[\bar{t}, \tau^{**}]$ relocate to $\bar{t}$, generating excess mass B (shaded) at the threshold and missing mass M (hatched) above it. Bunching is one-sided: Case-1 types ($\tau_L^* < \bar{t}$) do not relocate, so no compensating dip appears to the left of the bunching region.

market thickens.

(b) Timing. The interior optimum satisfies the logarithmic first-order condition $\Phi'(\tau^*)/\Phi(\tau^*) = r + \lambda_0$, which equates the proportional growth of the surgery surplus to the effective exit rate. A thicker market (higher λ_0) raises this hurdle; by the standard optimal-stopping logic—each additional instant of waiting now carries a larger probability of matching before investing—the optimum is pulled to earlier ages. Types whose smooth optimum was just past $\bar{t}$ shift to interior locations on $[t_0, \bar{t}]$, exiting Case 2 and the bunching set. This timing channel reinforces (a), which carries the formal weight.

Together, these forces shrink $\mathcal{B}$: $B(\rho)$ declines as the market thickens and vanishes in the thick-market limit. □

Intuition for Corollary 2. When the marriage market is thicker, matches arrive more quickly even without surgery. This lowers the incremental value of appearance investment and makes waiting less attractive. In thin or competitive female-side markets, by contrast, improving the matching hazard is more valuable, and more women have desired surgery ages near the threshold. The empirical counterpart in Section 5.3 is the stronger bunching in high-pressure marriage markets measured by female-to-male ratios on matchmaking platforms.

Corollary 3 (Substitution from non-beauty investment)**.** *Bunching is less pronounced in regions where women possess stronger non-beauty marital appeal.*

Proof of Corollary 3 (substitution from non-beauty appeal). The third comparative static asks how bunching changes when women have stronger non-beauty appeal. We show that the bunching indicator $\mathbf{1}[\theta \in \mathcal{B}]$ is weakly decreasing in θ under (A1)–(A3). As a result, any first-order stochastic shift in the type distribution to the right reduces the bunching mass.

Two θ-channels move the catchment, and both work against bunching:

- $\kappa'(\theta) \geq 0$ (A3): a higher κ raises the right-segment boost $\Delta_R = \kappa\Delta_L$ toward Δ_L, directly shrinking the gross-benefit gap $\mathcal{G}(\Delta_L) - \mathcal{G}(\Delta_R)$. Since κ does not enter the left-segment payoff $W(\bar{t})$ (where $\rho = 1$) but raises the right-segment payoff $W(\tau_R^*)$, a higher κ makes the woman strictly less likely to satisfy $W(\bar{t}) \geq W(\tau_R^*)$.

- $\phi'(\theta) < 0$ (A2): a lower ϕ shrinks the jump at $\bar{t}$ by the same mechanism as Corollary 1 (the jump J is increasing in the beauty channel, with $\partial J/\partial\phi = mg(\bar{t})[\mathcal{G}'(\Delta_L) - \kappa\mathcal{G}'(\Delta_R)] > 0$ under the regularity condition stated there), so a lower ϕ weakens the deadline.

The baseline hazard λ_0 is a market-level parameter and does not vary with θ; the baseline-matching channel of earlier formulations is therefore absent here. (If one additionally posited $\lambda_0'(\theta) \geq 0$—stronger non-beauty types also match faster—the surplus channel of Corollary 2 would reinforce the two channels above.) Under (A1)–(A3) at least one of the two channels is strict on a set of positive measure. Hence $\mathbf{1}[\theta \in \mathcal{B}]$ is weakly decreasing in θ and strictly decreasing on a positive-measure set.

Now let $F_H \succeq_{\text{FOSD}} F_L$ in the first-order stochastic-dominance order. For any weakly decreasing indicator $h(\theta)$, FOSD implies

$$\int h(\theta)\, dF_H(\theta) \leq \int h(\theta)\, dF_L(\theta).$$

Applying with $h(\theta) = \mathbf{1}[\theta \in \mathcal{B}]$ yields $B(\rho_H) \leq B(\rho_L)$, with strict inequality whenever $F_H \neq F_L$ and $\mathcal{B}$ has positive measure. □

Intuition for Corollary 3. The economic content is substitution. Women with stronger non-beauty appeal receive a smaller marginal benefit from cosmetic surgery and are less exposed to the reputational penalty. They therefore have weaker incentives to use beauty investment as a deadline response. The empirical counterparts in Section 5.4 are economic dependence, female educational attainment, and the local gender wage gap, all of which proxy for the strength of non-beauty alternatives. The predicted pattern is that bunching should concentrate where these alternatives are most constrained.

Corollary 4 (Placebo predictions)**.** *No bunching arises either (i) for procedures without marriage-market content, or (ii) among men, who do not face the same age-based norm.*

Proof of Corollary 4 (placebo predictions). The placebo results correspond to removing, in turn, one of the two ingredients needed for bunching.

(i) Functional placebo: $m = 0$. If a procedure has no marriage-market beauty content, then $\Delta \equiv 0$ from (5). From (7), $\Phi(\tau; \theta) = -c < 0$ for all τ, so surgery is never undertaken for marital reasons and there is no incentive to time the procedure around $\bar{t}$. The set $\mathcal{B} = \varnothing$, and no bunching arises.

(ii) Male placebo: $\kappa(\theta) \equiv 1$. If the norm penalty is absent, then $\rho \equiv 1$ on $[t_0, T]$, so Δ and hence Φ are continuous at $\bar{t}$. The jump in Lemma 2 vanishes ($\Phi_L = \Phi_R$), W is C^1 on $[t_0, T]$, and no type bunches: $\mathcal{B} = \varnothing$. $\square$

Intuition for Corollary 4. The placebo predictions clarify what the mechanism is, and what it is not. Functional procedures may occur for medical reasons, but their timing should not respond to a discontinuity in beauty returns; men may undertake cosmetic procedures, but they do not face the same age-based female stigma. The model therefore requires both a marriage-market return to beauty and a norm-induced discontinuity in that return. The absence of bunching for functional surgeries and for men, documented in Figures 1 and 2, is consistent with these two ingredients being jointly necessary.