The Gender Income Gap may not measure the Gender Welfare Gap: The Role of College Majors and Education Policy *

Charles Ka Yui Leung[†]

C. Y. Kelvin Yuen[‡]

April 9, 2026

Abstract

This paper develops a quantitative life-cycle model incorporating endogenous college major decisions to evaluate how education policies impact gender inequality. Quantitative analysis reveals important trade-offs between the income and welfare gaps. For example, we find that universal free college tuition exacerbates the gender income gap while simultaneously narrowing the gender welfare gap. Gender-based affirmative action policies and randomly assigning college majors have little impact on the gender income gap, yet substantially worsen gender welfare disparities. This tradeoff is largely absent in a benchmark without major choice, highlighting the role of within-college sorting across heterogeneous majors. These findings underscore the importance of jointly considering welfare outcomes when designing education policies aimed at gender equality.

JEL codes: H20, I20, J20

Keywords: college major as income-generating process; gender gap in income and welfare; tuition subsidy policy; efficiency; income inequality

*Acknowledgment: We are grateful to Eric French, Naijia Guo, Eric Hanushek, Lance Lochner, Jin Zhou, participants of City University Micro Workshop, for many helpful discussion, and the City University of Hong Kong, Hong Kong University of Science and Technology (Guangzhou) for financial support. The usual disclaimer applies.

[†]City University of Hong Kong; Email: kycleung@cityu.edu.hk

[‡]The Hong Kong University of Science and Technology (Guangzhou) and The Hong Kong University of Science and Technology; Email: kelvinyuen@ust.hk

1 Introduction

Persistent gender disparities in earnings continue to draw significant attention from economists, policymakers, and the public. In a recent comprehensive report on the gender gap, the OECD highlights that:

> "One issue is horizontal segregation, meaning that women and men tend to be concentrated in different sectors or jobs. Women tend to be overrepresented in fields that pay relatively lower wages, such as caregiving jobs, and underrepresented in fields with relatively higher wages, such as engineering jobs. Vertical segregation, meaning that men and women are concentrated in different job levels, also affects women's pay.... Governments have implemented a wide array of public policies in efforts to close the gender wage gap, including improving girls' and women's equal access to education; ... Yet even the most comprehensive policy approaches have not been enough to close the gender wage gap anywhere in the OECD." (OECD, 2023, pp. 13–16)[1]

Indeed, in the United States, although women have surpassed men in college enrollment and completion rates, significant gender differences remain in fields of study, particularly within science, technology, engineering, and mathematics (STEM), business, and health professions (Goldin, 2014; Blau and Kahn, 2017). These disparities in educational choices contribute directly to substantial and persistent gender wage gaps in the labor market (Altonji, Blom and Meghir, 2012). Given these trends, understanding how education policies influence gender disparities in income and welfare—especially through the lens of college major choices—is economically important and highly policy-relevant.

Existing research has extensively documented gender differences in earnings, college attendance, and field-of-study decisions (Goldin (2014), Goldin et al. (2017), among others). However, most structural analyses of educational policy typically abstract from detailed endogenous major choices and their gender-specific implications, implicitly treating higher education as a homogeneous good. Consequently, the interactions between education policy, field-of-study decisions, and gender inequality remain insufficiently understood. This gap is critical because educational decisions not

[1] Appendix A provides more discussion on the policy measures for gender equality among some OECD countries.

only influence income distributions but also shape individuals' lifetime welfare by interacting with heterogeneous preferences, abilities, and financial constraints.

This paper aims to address several key questions through the following steps. First, we develop a quantitative life-cycle model in which economic agents of different genders can choose whether to pursue a college degree (referred to as the *extensive margin*) and select different majors within college (referred to as the *intensive margin*). As in reality, *different college majors* in the model represent *distinct income-generating processes*, each associated with unique mean payoffs and risks. Additionally, economic agents may derive idiosyncratic utility from different college majors, making their choices a trade-off between income potential and personal enjoyment. We introduce a menu of realistic variations in per-student costs across majors and allow the government to offer differentiated subsidies for each college major. Another merit of our model is its incorporation of female fertility choices, allowing it to capture the potential interactions between educational and fertility decisions. We calibrate the model to match a comprehensive set of empirical facts using data from the National Longitudinal Survey of Youth 1979 (NLSY79). The parameter values are selected to ensure that the model accurately captures several stylized facts observed in the United States.[2] Our model enables us to *decompose* the gender gap and relate it to various factors, including initial wealth, accumulated experience, college choices, fertility, etc.

We then conduct a series of counterfactual policy experiments to compare labor market outcomes, social efficiency, and income inequality. Our analysis includes an examination of the gender gap in both income and utility. On the positive front, we find that the uniform tuition regime, where students in different college majors are charged the same tuition, is superior than other regimes in balancing the efficiency and equality.[3] Thus, the practice in the United States that the differences in tuition across college majors being minimal may indeed be the optimal policy choice.

On the normative front, we also find that at the current level of tuition subsidy, there may be a tradeoff between the gender income gap and gender utility gap. Any attempt to narrow the

[2]After circulating the first draft of this paper, we became aware of Gemici and Wiswall (2014), which also develops a life-cycle model incorporating college majors and genders. However, this paper differs from theirs in several key areas. First, their model is calibrated to match the Current Population Survey (CPS), which is a monthly and relatively short panel. In contrast, our model is calibrated to match the National Longitudinal Survey of Youth (NLSY), which is an annual and long-term panel. Second, this paper builds on the empirical work of Altonji and Zimmerman (2019), which provides per-student cost estimates across different college majors. This allows us to conduct counterfactual experiments that would otherwise be impossible and shifts our research focus in a significantly different direction. We also explicitly model the fertility choice and allows it to interact with the college decision.

[3]We will provide precise formula on how we measure and compare across various policy regimes.

gender pay gap through alternating the subsidy on college tuition could lead to a widening of gender utility gap, suggesting serious cautions in the pursue of reduction of gender income gap as a policy objective. For instance, we find that free tuition maximizes aggregate welfare but simultaneously widens the gender income gap due to differential responsiveness by gender in major selection. Conversely, gender-based affirmative action policies aimed at balancing representation across majors and randomly assigning majors have minimal impact on the gender income gap, yet significantly raising the gender welfare gap and lowering total welfare (large distributional effects), highlighting the crucial role of comparative advantage and preference alignment in educational choices.

To further elucidate these results, we demonstrate that explicitly modeling gender and college-major heterogeneity is crucial. Reduced models omitting major choices exhibits no tradeoff between the income and welfare gaps. Tuition policy primarily operates through the extensive margin of college attendance, and earnings gains for women translate more directly into welfare gains. This benchmark clarifies the economic mechanism behind the tradeoff and highlights the importance of modeling within-college choices when evaluating education policy. Our findings thus offer important policy insights, showing that the design of education policies targeting gender inequality requires careful consideration of both economic returns and welfare implications arising from individuals' heterogeneous preferences and constraints.

Like many works in economics, this paper builds on several strands of literature. For instance, in the gender gap literature, Barth, Kerr and Olivetti (2021) find that there are gender income gap within establishment and marriage plays an important role in explaining the gender income gap. Goldin, Kerr and Olivetti (2024) find that the "mother penalty" exist for women, and will diminish as the children grow up, but would not vanish. Erosa, Fuster and Restuccia (2016) also build a life-cycle model and calibrate to match the employment ratio and working hours by age across gender. Erosa et al. (2022) build a static Roy model for married individuals and study the gender differences. Bang (2021) find that the job flexibility of husband has a significant impact on the wife labor market participation and hours of work, and thus affect the gender income gap in equilibrium.

There is also a large literature on whether (and if so, in what form) of financial aid would enhance college enrollment and completion (Angrist, Autor and Pallais (2023), Black et al. (2023), Bulman et al. (2021), Caucutt and Lochner (2020), Lee and Seshadri (2019), among others). Apparently, the literature has not reached a consensus (Dynarski, Page and Judith Scott-Clayton (2022)). In this paper, we will show that the enrollment response to the government subsidy can be nonlinear.

Economists are aware of the importance of college majors (Altonji, Arcidiacono and Maurel (2016), Altonji, Kahn and Speer (2016), among others). For instance, Arcidiacono (2004) builds a 3-period model of college major choice and estimate the model with the NLS72 data. Andrews et al. (2024) focus on the major choices of the Texas students, Patnaik et al. (2022) focus on the New York University students. Choi et al. (2023) employ matched ACS and LEHD data and find that the earning of engineering and business majors grow faster than the humanities major. Some recent work employ cross-sectional data and study whether the college major choices are related to the family background (for instance, Leighton and Speer (2023)).

This paper is organized as follows. Section 2 introduces the data and highlights several stylized facts. Section 3 presents our model. We calibrate the model to the US economy in Section 4. Section 5 presents policy experiments and counterfactual analysis. The last section concludes. Additional details are reserved in the Appendix.

2 Empirical Facts

Our empirical analysis draws primarily on data from the National Longitudinal Survey of Youth 1979 (NLSY79), a comprehensive panel dataset covering a nationally representative sample of individuals in the United States. The NLSY79 tracks detailed demographic information, educational choices, labor market experiences, and earnings, allowing us to capture life-cycle dynamics accurately. The NLSY79 data remain nationally representative for the cohort studied, enabling robust and generalizable conclusions about life-cycle behavior and policy impacts (Bick, Blandin and Rogerson, 2025). We use the period from 1979 to 2016. The final sample contains 181,830 person-year observations with consistent longitudinal information on demographics, education, employment, and income.[4]

We leverage extensive earnings data available from the NLSY79, which captures respondents' income from wages on an annual basis. Earnings data are consistently reported, allowing us to

[4] See Appendix B for detailed data cleaning process.

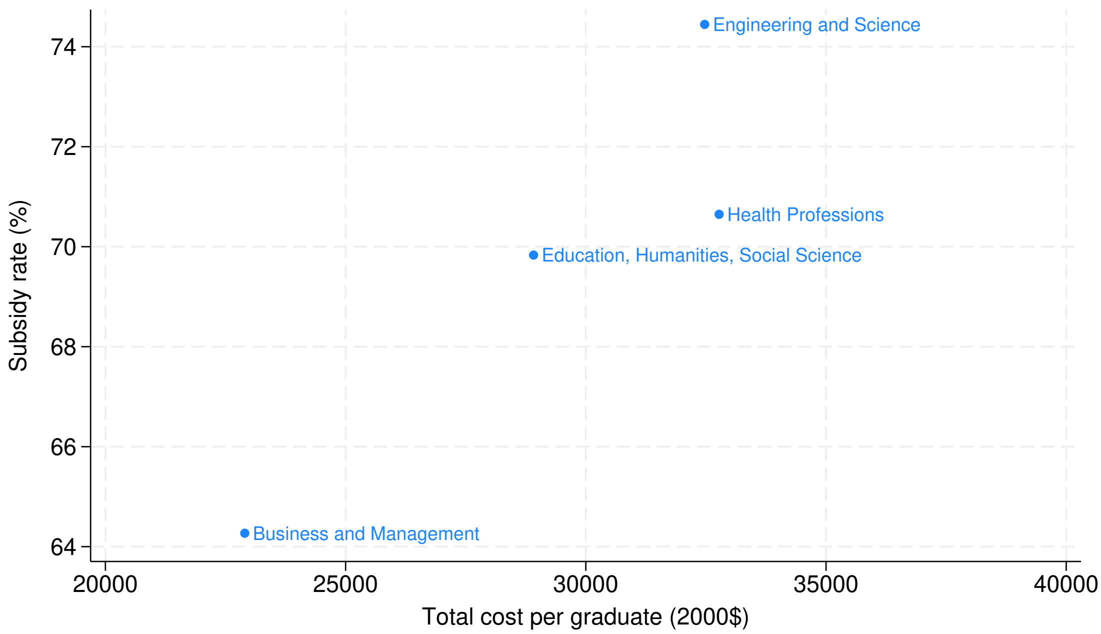

Figure 1: Subsidy rate vs. tuition by major

Source: Altonji and Zimmerman (2019) and authors' calculation.

construct precise life-cycle earnings profiles differentiated by gender and major. All earnings figures are deflated to 2000 dollars. Furthermore, information on hours worked enables us to construct reliable measures of hourly wages, a critical component in accurately assessing gender income gaps across majors.

Our model highlights the differences across various college majors. Based on Altonji and Zimmerman (2019), we aggregate various college majors into four main categories (Business and Management, Education, Humanities, Social Sciences, Engineering & Science, and Health Professions) to calculate the average cost, tuition, and corresponding subsidies on a per-student basis.[5] Combining their data with the NLSY79, several stylized facts emerge naturally.

Stylized fact 1: Figure 1 illustrates significant differences in total costs per student across college majors. While tuition remains relatively uniform, subsidy rates—defined as the ratio of the difference between cost and tuition to total cost—vary drastically.

[5] An exception is the law major, which has a distinctly different income trajectory over the life cycle. As a result, it cannot be classified within any of the major categories. Additionally, its student population share is minimal, making it impractical to establish it as a fifth college major group. Consequently, we have excluded the law major from our analysis. Table B.1 provides the details of the classification.

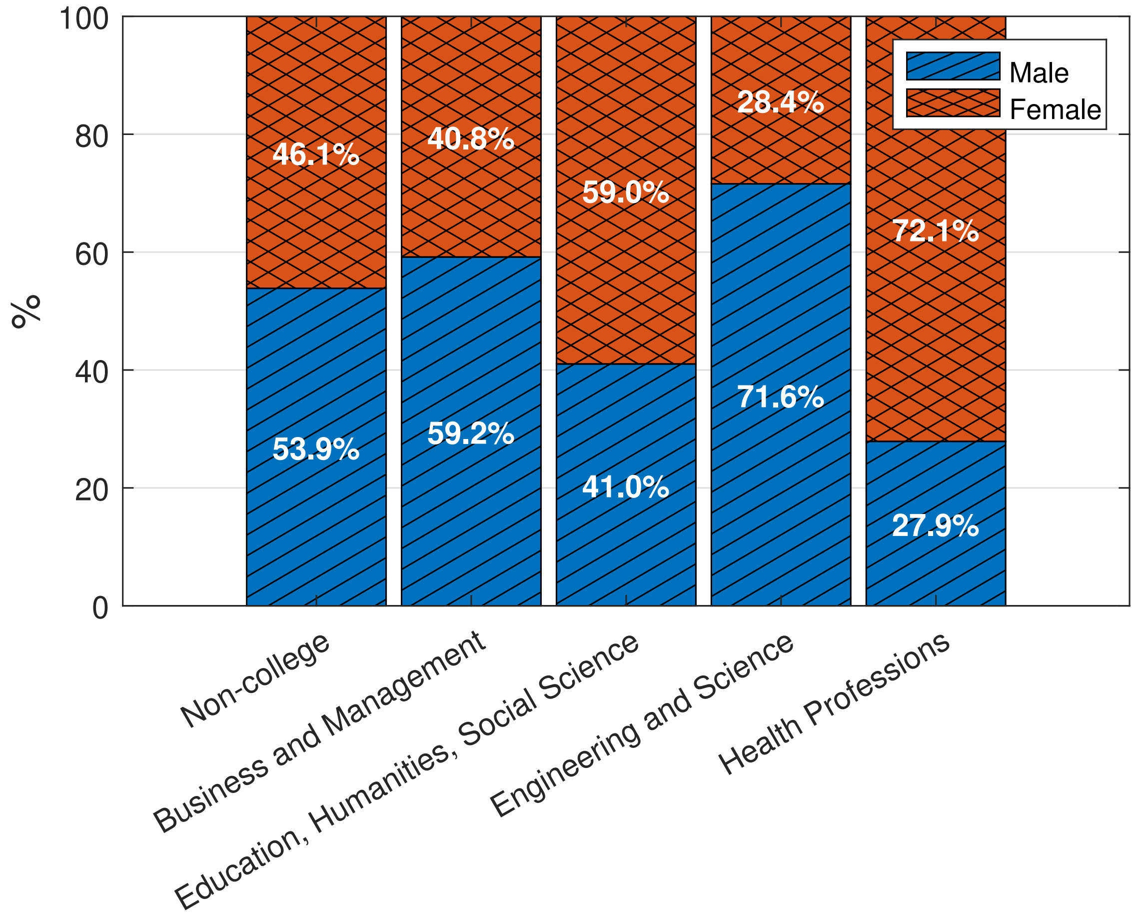

Figure 2: Gender composition by major

Source: The National Longitudinal Survey of Youth 1979 (NLSY79), 1979-2016.

Stylized fact 2: College majors differ notably in gender composition, as shown by Figure 2. Specifically, approximately 71% of students in Engineering and Science are male, while about 72% of students in Health fields are female.[6]

Stylized fact 3: Earning profiles by college major *diverge* rather than converge throughout the life cycle. Figure 3 illustrates the mean income profiles across different majors, including individuals who did not attend college. For the NLSY79 cohort, those majoring in Business and Management tend to earn more than their peers in other fields. Majors in Engineering and Science have earnings that are approximately equal to those in Health Professions. In contrast, students majoring in Education, Humanities, and Social Sciences earn more than those who did not attend college, but they still fall short compared to other majors.[7]

[6] Our results are broadly consistent with the existing literature, such as Goldin (2014), Gottlieb et al. (2025).

[7] Appendix B.2 formally rejects the hypothesis that different college majors share the same income generating

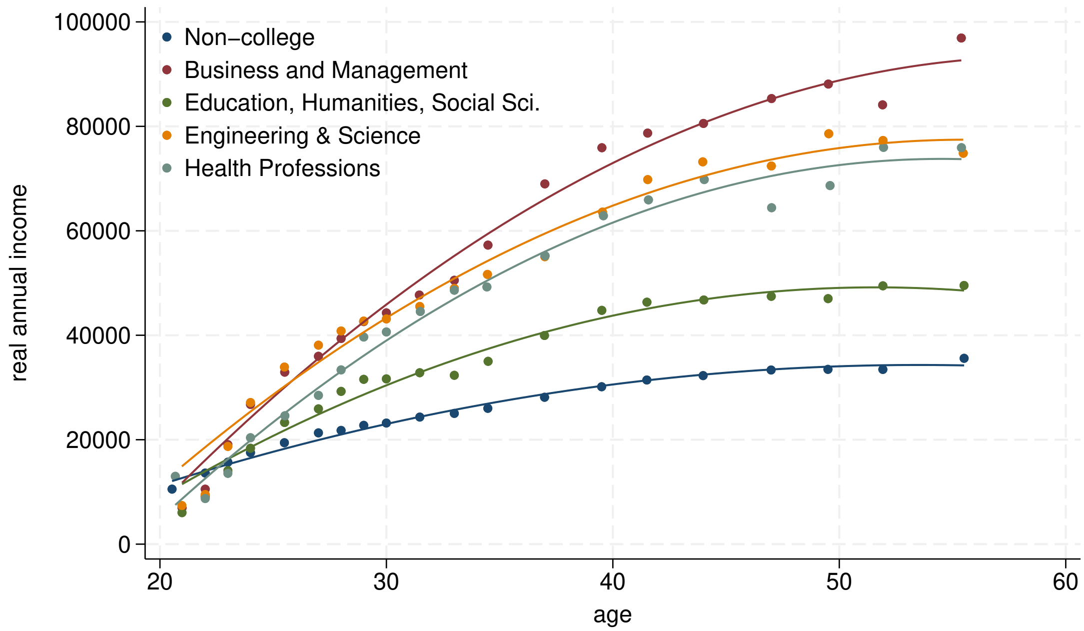

Figure 3: Life-cycle earnings by major

Source: The National Longitudinal Survey of Youth 1979 (NLSY79), 1979-2016.

Stylized fact 4: The gender income gap is *not* a constant; it fluctuates over the life cycle and varies by college major. Figure 4 shows that, for non-college attendees, the gap rises slightly between ages 20 and 40 before leveling off or diminishing. However, college graduates experience *different* trends. Between ages 20 and 30, the gender gap in Health Professions is relatively low, similar to other majors and below that of non-attendees. After age 30, the gap in Health Professions increasingly surpasses those in other fields. Conversely, the gap in Education, Humanities, and Social Sciences is more aligned with non-college goers, while the gaps in Business, Management, Engineering, and Science continue to widen with age.

These graphs inspire several questions. First, given the dramatically different income trajectories and costs per student associated with various college major groups, should we subsidize the higher education?[8] Second, even if we accept that government subsidies for college students are necessary, the most socially desirable form of subsidy remains uncertain. For example, should we maintain the current system where each student pays the same tuition, or should all students re-

process.

[8]See section C for more discussion.

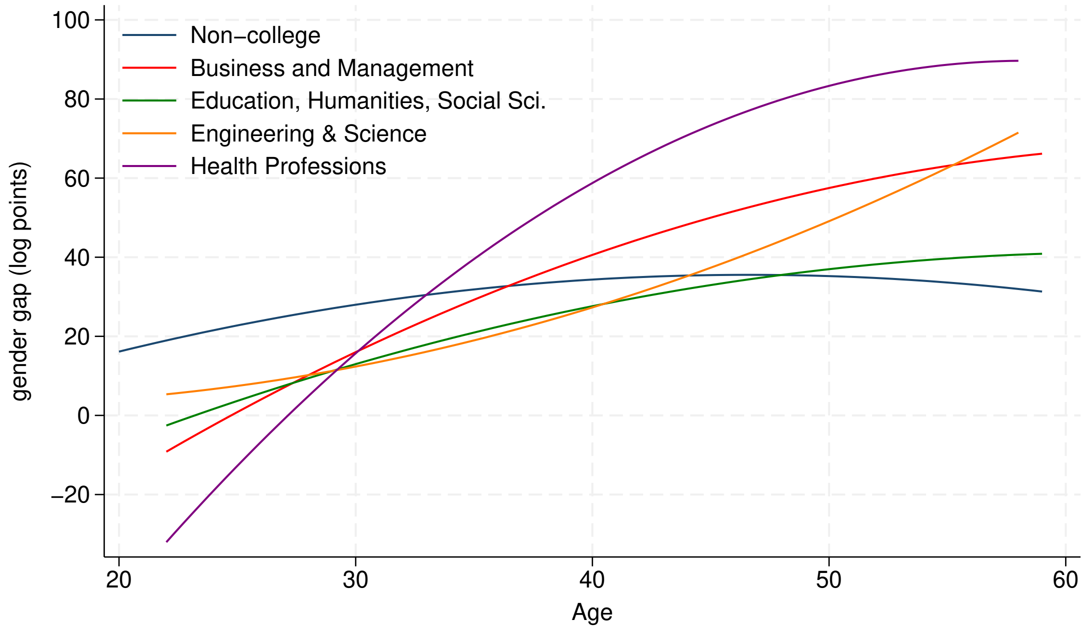

Figure 4: Life-cycle gender income gap by major

Source: The National Longitudinal Survey of Youth 1979 (NLSY79), 1979-2016.

ceive equal subsidies? Should the student subsidies be identical across majors in absolute terms, or should the subsidy rate be equalized across majors? Previous research has demonstrated that government aid can influence students' choice of major Sjoquist and Winters (2015). Does this governmental support also impact the gender income gap? To address these questions, we develop a life-cycle model that incorporates explicit college major choices, accounting for differences in costs, income processes, and consumption value.[9]

3 Model

Following the spirit of Keane and Wolpin (1997), we develop a life-cycle model that incorporates the choice of major college.[10]

[9]The consumption value of college has been established in several studies, see Gong et al. (2018), Guo and Leung (2021), Jacob, McCall and Stange (2018), among others.

[10]Some authors have explored the importance about the beliefs or expectations on the college major choices and subsequent labor market performance (Arcidiacono, Hotz and Kang (2012), Stinebrickner and Stinebrickner (2014)). For simplicity, we focus on the case where agents have rational expectation and leave the modeling of beliefs and expectations to future research.

3.1 Environment

The economy consists of N individuals, which are categorized by gender type $s \in \{n, f\}$, where n denotes males and f females. Females make the same choices as males; however, they also need to consider the fertility choice.[11] Individuals are risk averse with common discount factor β and a Constant Relative Risk Aversion (CRRA) utility function:

$$u(c) = \frac{c^{1-\sigma} - 1}{1 - \sigma} \tag{1}$$

where σ measures the coefficient of relative risk aversion. Individuals' age is indexed by $t \in \{0, 1, 2, \dots, T\}$ up to a maximum age T. There is no direct utility from leisure.

There are M college majors, with college major types indicated by $m \in \{0, 1, 2, \dots, M\}$, where $m = 0$ indicates no college enrollment. Individuals are heterogeneous in their ability, captured by a vector $A = \{A_0, A_m\}_m$, comprising general ability A_0 and major-specific abilities A_m. These abilities follow normal distributions conditional on gender type, $A_{s,m} \sim N(\mu_A^{s,m}, \sigma_A^{s,m})$. Additionally, each individual receives flow utility $b_{s,m}$ from majors during school, drawn from a normal distribution with mean zero conditional on gender type and major, $b_{s,m} \sim N(0, \sigma_b^{s,m})$. Finally, individuals hold assets k, which evolve through saving and borrowing decisions over their life cycle.

Upon choosing a college major, individuals decide whether to finance their education through student loans, indicated by $d \in \{0, 1\}$, with $d = 1$ representing borrowing. The total student loan amount borrowed is denoted by $\bar{S}$. Following Matsuda and Mazur (2022), repayments occur during the working stage under one of two schemes: Fixed Repayment (FR) or Income-Contingent Repayment (ICR). Under FR, the annual repayment P_t^{FR} is given by:

$$P_t^{FR} = \frac{r\bar{S}}{1 - (1 + r)^{-T^{FR}}} \tag{2}$$

where r is the risk-free interest rate and T^{FR} is the period of repayment.

Under the ICR scheme, repayments depend on income y relative to a threshold $\bar{y}$, formulated as:

$$P_t^{ICR} = \min\{\omega \max\{y - \bar{y}, 0\}, P_t^{FR}\} \tag{3}$$

[11] Empirically, the labor supply and labor market performances are drastically different between females with and without children (Barth, Kerr and Olivetti, 2021; Goldin, Kerr and Olivetti, 2024).

Thus, repayment schedules significantly shape individual decisions on college major selection and subsequent financial trajectories.[12]

3.2 Working Stage

Labor productivity in the model is determined by a human capital index that reflects both age and accumulated work experience, the latter measured in cumulative hours worked. Formally, the cumulative hours of experience up to period t as

$$Exp_{s,m,t} = \sum_{\tau=1}^{t} h_{s,m,\tau}$$

where $h_{s,m,\tau}$ is the hours worked in period t. The human capital index is a polynomial function of age and experience:

$$\begin{aligned} logH_{s,m,t} = {} & \beta_{s,m,1} Exp_{s,m,t} + \beta_{s,m,2} Exp^2_{s,m,t} + \beta_{s,m,3} Exp^3_{s,m,t} + \beta_{s,m,4} Age_t + \beta_{s,m,5} Age^2_t \\ & + \beta_{s,m,6} Age^3_t \end{aligned}$$

For females, human capital depends also on fertility timing. We estimate separate human capital functions for females with no child and females with child, and splice them together according to the chosen age at first birth.[13]

Given the human capital function, the hourly wage $w_{s,m,t}$ is specified as

$$w_{s,m,t} = w_{s0} H_{s,m,t} \exp(\omega_{s0} A_0 + \omega_{sm} A_m + \varepsilon_t) \tag{4}$$

where w_{s0} represents the baseline productivity level for gender s, and the idiosyncratic earnings shock $\varepsilon_t \sim N\left(0, \sigma_\varepsilon^{s,m}\right)$. Hence, the annual income is then given by

$$y_{s,m,t} = w_{s,m,t} h_{s,m,t}$$

[12] Following Matsuda and Mazur (2022), we abstract from student loan default decision. In the US, student loans typically cannot be discharged under Chapter 13 bankruptcy proceedings, significantly limiting the economic incentives and feasibility of defaulting (Darolia and Ritter, 2020).

[13] Specifically, before the child arrives, the relevant coefficients $\beta_{s,m,k}$ and hours profile are those for females with no child; after birth, the coefficients and hours switch to those for females with child.

The model incorporates a progressive income tax schedule $\tilde{\tau}(y)$, defined as:

$$\tilde{\tau}(y) = \max\{y - \theta y^{1-\tau}, 0\}, \tag{5}$$

where τ and θ are parameters governing, respectively, the progressivity and scale of the tax system.[14]

The individual optimization problem at age t, conditional on gender type s, college major m, initial asset k_t, abilities A, idiosyncratic income shock $\varepsilon_{s,m}$, and loan decision d, is hence formulated by the Bellman equation:

$$W_{s,m,t}(k_t, A, \varepsilon_t, d) = \max_{c_t, k_{t+1}} \{u(c_t) + \beta E_\varepsilon [W_{s,m,t+1}(k_{t+1}, A, \varepsilon_{t+1}, d)]\}, \tag{6}$$

subject to the budget constraint:

$$c_t + k_{t+1} + dP_t = k_t(1+r) + y_{s,m,t} - \tau(y_{s,m,t}). \tag{7}$$

Here, dP_t is the student loan repayment amount at period t, conditional on the repayment scheme chosen earlier.

3.3 Schooling Stage

At the beginning of the schooling stage, each individual is endowed with initial wealth k_0 and a general ability level A_0. The joint distribution of the initial conditions, A_0 and the logarithm of initial wealth $\log(k_0)$, follows a bivariate normal distribution, with mean $\mu = (0, 0)$ and covariance matrix Σ_s, conditional on gender type s:

$$\Sigma_s = \begin{pmatrix} \left(\sigma^s_{k_0}\right)^2 & \rho^s_{k_0,A_0}\sigma^s_{A_0}\sigma^s_{k_0} \\ \rho^s_{k_0,A_0}\sigma^s_{A_0}\sigma^s_{k_0} & \left(\sigma^s_{A_0}\right)^2 \end{pmatrix} \tag{8}$$

When choosing to enroll in college, an individual assesses the value of selecting major m at a four-year institution. Financial aid for a student with ability percentile $q(A_0)$ and initial wealth k_0 is given by

$$Aid_m(A_0, k_0) = \alpha_{m0} + \alpha_{m1}q(A_0) + \alpha_{m2}q(A_0)^2 + \alpha_{m3}k_0$$

[14]The formulation of the progressive income tax schedule follows Guner, Lopez-Daneri and Ventura (2016).

The individual's value function for choosing major m is given by:

$$V_{s,m}(k_0, b, A) = \max_{c_0, k_4, d \in \{0,1\}} \left\{ u(c_0) + b_{s,m} + \beta^4 E_\varepsilon [W_{s,m,4}(k_4, A, \varepsilon_{s,m}, d)] \right\} \tag{9}$$

subject to the budget constraint:

$$c_0 + k_4 = k_0(1+r)^4 + Aid_m + d\bar{S} - T_m, \tag{10}$$

where consumption c_0 during schooling, $k_4 \geq 0$ is the asset carried forward after four years of college, T_m is the (net) tuition cost associated with major m. The major-specific non-pecuniary benefit or taste $b_{s,m}$ is drawn from a normal distribution, $b_{s,m} \sim N(0, \sigma_b^{s,m})$.

The optimal choice of college major or opting out of college is based on a comparison of values across available alternatives. There is a cutoff general ability $\bar{A_0^m}$ to enroll in each major, which is determined using a Gale–Shapley algorithm (Gale and Shapley, 1962). Subject to the cutoff scores, the individual's life-time utility can be written as

$$V(k_0, b, A) = \max_{m: A_0 \geq \bar{A_0^m}} \{ V_{s,m}(k_0, b, A), E_\varepsilon [W_{s,m,0}(k_0, A, \varepsilon, 0)] \}, \tag{11}$$

where the second term inside braces captures the expected value of entering directly into the workforce without attending college. Thus, individuals choose between enrollment in a particular college major and immediate labor market participation based on maximizing expected lifetime utility.

3.4 Fertility Decisions

Women determine whether and when to have child during their working stage. Let $t_f \in \{0, 1, ..., T_f + 1\}$ be the fertility time, where T_f is the maximum fertility age and $t_f = T_f + 1$ corresponds to "never" (no child). There is a one-time fertility utility benefit $b_f \in \{0, \bar{b}_f\}$ where $Pr(b_f = \bar{b}_f) = p_f$ and $Pr(b_f = 0) = 1 - p_f$. Each woman chooses her optimal fertility timing conditional on her fertility benefit. Specifically, for a particular fertility timing t_f, the value of working at $t_f - 1$ is then given by

$$W_{f,m,t_f-1}(k_t, A, \varepsilon_t, d) = \max_{c_t, k_{t+1}} \left\{ u(c_t) + \beta E_\varepsilon \left[W_{f,m,t_f}(k_{t+1}, A, \varepsilon_{t+1}, d) + b_f \right] \right\}$$

The problem is solved for all t_f. A woman with benefits b_f will then choose the optimal timing t_f^* with the highest life-time utility.

3.5 Government Budget Constraint

The government balances the budget by equating total revenues and total expenditures associated with education financing and taxation over the life-cycle of individuals. The budget constraint can be expressed as:

$$\sum_{t=1}^{T} \sum_{i=1}^{N} \tau(y_{it}) + \sum_{m=1}^{M} \sum_{i \in \text{major } m} T_m + \sum_{i:d_i=1} \sum_{t} P_t = \sum_{m=1}^{M} \sum_{i \in \text{major } m} K_m + \sum_{i:d_i=1} \sum_{t=1}^{T} \bar{S}. \tag{12}$$

On the revenue side, the government collects resources from three primary sources. First, total tax receipts $\sum_{t=1}^{T} \sum_{i=1}^{N} \tau(y_{it})$ aggregate individual income taxes paid over the lifetime of each agent. Second, $\sum_{m=1}^{M} \sum_{i \in \text{major } m} T_m$ denotes the total tuition payments received from students enrolled in each major. Third, $\sum_{i:d_i=1} \sum_t P_t$ represents repayments collected from student loan borrowers under their respective repayment schemes.

On the expenditure side, the government incurs two primary costs. First, $\sum_{m=1}^{M} \sum_{i \in \text{major } m} K_m$ reflects the total education costs associated with each major, including subsidies, operational costs, faculty salaries, infrastructure, and other resources required for delivering education (Altonji and Zimmerman, 2019). Second, $\sum_{i:d_i=1} \sum_{t=1}^{T} \bar{S}$ indicates the total amount of student loans provided to students who choose to finance their education through borrowing.

3.6 Gender Gaps and Welfare

In evaluating the outcomes of various educational policies, it is crucial to distinguish between the gender income gap (GIG) and the gender welfare gap (GWG). Let the population size of each gender be denoted by $\{N_n, N_f\}$. Suppose individual i's chosen college major is $m(i)$ and their gender is $s(i)$. The gender income gap for the economy is defined as the difference in the average log hourly wage between males and females:

$$GenderGap_{Income} = \frac{1}{N_n} \sum_{i:s(i)=n} \log(w_{it}) - \frac{1}{N_f} \sum_{i:s(i)=f} \log(w_{it}) \tag{13}$$

where w_{it} represents individual i's hourly wage at time $t$given by (4). We can define the GIG conditioning on enrolling in the same college major analogously.

By the same token, we measure the aggregate gender welfare gap as the difference in the average lifetime utility between males and females:

$$GenderGap_{Welfare} = \frac{1}{N_n} \sum_{i:s(i)=n} V_i - \frac{1}{N_f} \sum_{i:s(i)=f} V_i \tag{14}$$

where V_i denotes the lifetime utility of individual i. And we can define the GWG conditioning on enrolling in the same college major analogously.

Finally, overall welfare is computed according to the utilitarian criterion, aggregating the lifetime utilities of all individuals:

$$Welfare = \sum_{i=1}^{N} V_i \tag{15}$$

One of our major results stems from the differentiation between different measures of gender disparity, GIG and GWG. The latter explicitly recognize both pecuniary and non-pecuniary dimensions of individual welfare. In the next sections, we will quantitatively examine the implications of this distinction in evaluating the impacts of various education policies.

4 Quantitative Analysis

In this section, we estimate the model using the NLSY79 data. The model is able to replicate life-cycle earnings profile and gender income gap across various college majors in the data.

4.1 Calibration and Identification

Following the literature, we calibrate the model in two steps (Keane, Todd and Wolpin (2011)). First, we impose parameter values on the basis of the data and other studies. We estimate from the NLSY79 all objects that can be directly observable: human capital profiles by major and sex, hours

Table 1: Calibration parameters

Variable	Description	Target/source
T	Length of working life	40 Years of working life
r	Interest rate	5%
β	Discount factor	0.96; discount rate of 4%
σ	Relative risk aversion	2
T_f	Maximum fertility age	35
w	Wage rate	Normalize average wage of non-college workers at 20
$\sigma^s_{k_0}$	Initial wealth distribution	Var(wealth at age 30)
$\sigma^{s,m}_\varepsilon$	S.D. of income shock	2-year persistence of income
$\{\alpha^{s,m}_1, \alpha^{s,m}_2\}$	Learning profile	Average income at age 30 and 50
T_m	Tuition fee	NLSY79 Data
$\mu^{s,m}_A$	Mean ability	Average wage income
$\sigma^{s,m}_A$	S.D. of ability	Variance of income at age 30
$\sigma^{s,m}_b$	S.D. of taste shock	Average ability of majors
p_f	Probability of high-fertility type	Fertility rate
$\bar{b}_f$	Fertility benefits (high)	Average age at first birth
$\rho^s_{k_0,A_0}$	Correlation between k_0 and A_0	NLSY79 Data
$\underline{A}^{s,m}_0$	Cutoff ability for majors	Share of majors
$\bar{S}$	Student loan	Average amount of student loan
τ	Progressivity of income tax	Guner, Lopez-Daneri and Ventura (2016)
θ	Scaling parameter of income tax	Balanced budget of the government

profiles, the variance of idiosyncratic wage shocks, tuition and costs, the financial aid schedule, and a small set of preference and policy parameters (discount factor, interest rate, tax curvature, loan terms). In a second step we calibrate male specific parameters and female specific parameters by method of simulated moments. We summarize the calibration parameters in Table 1. Most of the assigned parameters are compatible with the literature.

We define one model period as one year, with individuals beginning the life cycle at age 20 and facing a time horizon of $T = 40$. As outlined in Section 2, our analysis includes four distinct college major categories: Business and Management; Education, Humanities, and Social Science; Engineering and Science; and Health Professions.

The risk-free interest rate r is set at 5%. The subjective discount factor β is assumed to be 0.96, reflecting an annual discount rate of approximately 4%. The relative risk aversion coefficient σ is fixed at 2, aligning with empirical estimates derived from micro-level consumption data (Attanasio and Weber, 1995).

The correlation between initial wealth and general ability $\rho^s_{k_0,A_0}$ is calculated using data from the NLSY79, employing Armed Forces Qualification Test (AFQT) scores as proxies for individual ability. Average student loan amounts $\bar{S}$ are also derived from the NLSY79, considering only individuals who held a positive student loan balance at age 24. Tuition fees T_m for the four majors are similarly computed from NLSY79 data, averaging approximately \$2,000, \$2,200, \$2,100, and \$2,400 per year (in 2000 dollars) respectively. Educational costs are taken from Altonji and Zimmerman (2019), who derive these values based on data from the Board of Governors of the Florida State University System (FLBOG2000-2014). The annual hours worked $h_{s,m,t}$ is measured directly from the NLSY79 data.

The earnings parameters $\{\beta_{s,m,k}, \sigma_{\varepsilon}^{s,m}\}$ are estimated by running a regression of log-wage on experience and age with individual fixed effects for each $\{s, m, t_f\}$ combination. The financial aid coefficients $\{\alpha_{m,k}\}$ are are estimated from aid regression for each major. See Table B.3 for the financial aid regression results, Table B.4 for the estimated wage regression coefficients, and Table B.5 for the estimated income shocks dispersion.

The remaining parameters are calibrated jointly by matching a set of targeted moments in the data. First, baseline productivity w_s is calibrated by targeting the average income of non-college workers at age 20. Also, the means of the ability $\{\mu_{A_m}^s\}$ are identified by the cross major pattern of mean log earnings, in particular the college premia relative to non college. the dispersions of ability $\{\sigma_{A_m}^s\}$ control the amount of permanent cross sectional heterogeneity in earnings beyond the idiosyncratic shocks, which are pinned down by the variance of log earnings at age 30 by major. The standard deviation of the initial wealth distribution, $\sigma_{k_0}^s$, is calibrated by matching observed variance in wealth at age 30. The standard deviation of taste shocks, $\sigma_b^{s,m}$, is determined by aligning model-generated average ability (in percentiles) within each major to the observed AFQT scores. Finally, for women, the fertility parameters $\{p_f, \bar{b_f}\}$ together determine both how many women have children and when they choose to have them. They are identified by matching the overall fertility rate and the average age at first birth.

We employ the Gale-Shapley algorithm (Gale and Shapley, 1962) to determine the cutoff ability $\bar{A}_0$ required for enrollment in each major, calibrating this to match observed enrollment shares in the NLSY79 dataset. Finally, the progressivity parameter of the income tax system is set at 0.053 following Guner, Lopez-Daneri and Ventura (2016), and the scaling parameter θ of the tax system is calibrated to ensure that the government's budget constraint holds in the baseline as well as in all policy experiments.

Table 2: Targeted moments: data vs. model

Moment	Men Data	Men Model	Women Data	Women Model
Mean log annual earnings, by major				
Non-college	10.0750	9.9572	9.5537	9.4669
Business & Management	10.8721	10.8765	10.3426	10.1014
Educ./Hum./Soc. Sci.	10.4351	10.4350	10.0278	9.8527
Engineering & Science	10.6903	10.6898	10.2699	10.0090
Health Professions	10.8691	10.9346	10.2993	10.0408
Var(log earnings) at age 30, by major				
Non-college	0.6520	0.6861	1.1047	1.2060
Business & Management	0.3789	0.3688	0.7550	0.7238
Educ./Hum./Soc. Sci.	0.4479	0.4390	1.2615	1.2628
Engineering & Science	0.4479	0.4521	0.3972	0.3803
Health Professions	0.5312	0.5165	0.5629	0.5853
Mean AFQT percentile by college major				
Business & Management	76.6615	72.7069	72.6685	72.3928
Educ./Hum./Soc. Sci.	72.0371	73.0900	69.6736	64.4990
Engineering & Science	79.0283	75.5354	76.3130	73.4084
Health Professions	78.4738	77.8435	68.5646	65.4227
Wealth and fertility				
Var(wealth at age 30)	3.21×10^{10}	3.32×10^{10}	2.50×10^{10}	2.59×10^{10}
Fertility rate (ever had child)	–	–	0.8077	0.7750
Mean age at first birth	–	–	24.0	23.0191

To account for potential heterogeneity between different genders, we select the parameter vector Θ^s for each gender s to minimize the squared proportional distance between the model-implied moments $m^s\hat{(\Theta^s)}$ and the empirical moments m^s:

$$\Theta^s = arg \min_{\Theta} \left\| W^{1/2}(m^s\hat{(\Theta^s)} - m^s) \right\|$$

where W is a diagonal weight matrix which is proportional to squared percentage deviation for each moment. The objective is evaluated by simulating a cohort of $N = 10,000$ individuals from the joint distribution of (k_0, A_0, A_m, b_m, B), solving the dynamic program, running the college admissions mechanism, and computing simulated moments from the resulting life cycle paths.

Table 2 compares the targeted empirical moments with their corresponding model-generated counterparts. The model generally achieves a good fit to the data, validating the calibration

Table 3: Estimated values of parameters

Block	Parameter	Men	Women
Wage level and initial wealth			
	w (wage level scale at age 30)	4.9471	3.7969
	σ_{k0} (s.d. of initial wealth)	23.1246	15.2461
Ability means by major			
	$\mu_{A,1}$ (non-college; normalized)	0	0
	$\mu_{A,2}$ (Business & Management)	-0.1870	0.6772
	$\mu_{A,3}$ (Educ./Hum./Soc. Sci.)	-0.5125	1.0438
	$\mu_{A,4}$ (Engineering & Science)	-0.3157	-5.4906
	$\mu_{A,5}$ (Health Professions)	-0.3735	-3.5722
Ability dispersion by major			
	$\sigma_{A,1}$ (non-college)	0.6056	0.8500
	$\sigma_{A,2}$ (Business & Management)	0.1500	0.6001
	$\sigma_{A,3}$ (Educ./Hum./Soc. Sci.)	0.3374	0.9014
	$\sigma_{A,4}$ (Engineering & Science)	0.1379	0.3168
	$\sigma_{A,5}$ (Health Professions)	0.2980	0.5213
Taste dispersion for college majors			
	$\sigma_{b,2}$ (Business & Management)	6.5512	201.0000
	$\sigma_{b,3}$ (Educ./Hum./Soc. Sci.)	6.6165	505.0000
	$\sigma_{b,4}$ (Engineering & Science)	6.5462	75.1200
	$\sigma_{b,5}$ (Health Professions)	2.6024	75.1200
Fertility parameters (women only)			
	b_f^H (fertility utility benefit for high type)	–	11.4685
	p_f^H (probability of high fertility type)	–	0.8077

Notes: Major indexing follows the model convention: $m = 1$ is non-college and $m = 2, \ldots, 5$ are the four college majors listed in the table. The normalization $\mu_{A,1} = 0$ is imposed in both calibrations. In the female model code, the low-type fertility benefit is normalized to $b_f^L = 0$.

strategy.

Table 3 presents the estimated values of parameters. Some parameters display significant variations across gender.[15] For example, the taste dispersion for college majors (measured by the standard deviations of the taste shock) among males never exceeds seven across different fields. In contrast, the dispersion for females is at least ten times greater: approximately 75 for Engineering and Science, around 200 for Business and Management, and reaching 500 for Education, Humanities, and Social Sciences. This stronger non-pecuniary preference among females suggests

[15] To enhance robustness, we also "winsorize" by excluding the top and bottom 1% of taste shock realizations. The results, available upon request, remain essentially unchanged. This indicates that our findings are not driven by a few extreme values at either end of the distribution.

that some may be 'price-insensitive,' meaning their college choices may remain unchanged despite tuition subsidies or varying labor market rewards across majors.

The differences between genders also extend to the distribution of abilities, although not as pronounced as in taste. First, we normalize the means of non-college outcomes for both genders to zero. We find that the ability means for males are all *negative* across different college majors. In contrast, females show positive ability means in Business and Management, as well as in Education, Humanities, and Social Sciences, indicating that females have an absolute advantage in these fields over males.

Additionally, the ability dispersion, measured by the standard deviation of ability, varies between genders. For instance, while $\sigma_{A,2}$ is estimated at 0.15 for males, it is 0.60 for females—almost four times as large. Similarly, $\sigma_{A,3}$ is approximately 0.34 for males, but rises to 0.90 for females, nearly three times as large. For $\sigma_{A,4}$, the values are around 0.14 for males compared to 0.32 for females, more than double. Finally, $\sigma_{A,5}$ shows 0.30 for males and 0.52 for females, almost twice as large. Given these differences in ability, it is only natural that males and females choose college majors differently.

Finally, initial wealth dispersion (σ_{k_0}) exhibits significant differences between genders: it is higher among males (23.12), and lower for females (15.25). This suggests meaningful disparities in initial economic conditions that may influence educational choices and subsequent career trajectories.

Overall, these parameter estimates highlight significant heterogeneity in tastes, abilities, productivity, and initial conditions across genders and fields of study, providing valuable insights into the sources and nature of observed gender disparities in educational and labor market outcomes.

4.2 Life-cycle profiles and Gender gap

Figure 5 compares life-cycle earnings profiles by gender and college major choices. While the data is generally noisier than the model predictions, the model effectively captures key patterns, especially the rapid earnings growth in younger ages, peak earnings around ages 40–50, and relative flattening or decline thereafter [16].

[16] Figure E.1 in the Appendix shows that the model matches the life-cycle income profiles well for females with different fertility status.

Figure 6 illustrates the model's fit regarding the gender income gap, defined as the average difference in log hourly wages between males and females, across different fields of study. The model predictions closely match the observed data, effectively capturing the magnitude and variation of gender income disparities across majors. There are notable variations in the gender gap across majors. The largest gender gaps are observed in Business and Management and Health Professions majors, where the gap exceeds 30%. These results reflect substantial gender differences in earnings potential within these relatively high-paying fields between males and females. In contrast, Education, Humanities, and Social Science and Engineering & Science majors show relatively smaller gender gaps at approximately 20%, suggesting lower gender disparities in earnings within these fields. This smaller gap could reflect more similar labor market attachment, hours worked, or career trajectories for men and women in these areas, which typically offer more flexible or family-friendly employment arrangements. Overall, the model successfully replicates these important stylized facts, confirming its suitability for evaluating educational policies.

4.3 Decomposition of the Gender Gap

The estimated structural models offer the advantage of decomposing the gender gap, allowing us to attribute contributions to each factor. To attribute the gender gap to distinct channels, we conduct a sequence of counterfactual exercises that progressively replace female-side objects and parameters with their calibrated female values, starting from a benchmark in which women face the male economy. Concretely, we construct a baseline counterfactual economy in which women are assigned (i) the male productivity scaling, (ii) the male initial wealth distribution, (iii) the male ability distribution, (iv) the male major-specific taste dispersion, (v) male admission cutoffs, (vi) male hours profiles, (vii) male human capital profiles, and (viii) male income shock processes. Fertility is shut down in this baseline by setting the female fertility utility parameter and the share of high-fertility-preference types to zero. In this counterfactural baseline, the model delivers (approximately) zero gender gap by construction. We then introduce each female-specific component in the above order.

Figure 7 provides a visualization.[17] The model implies a total gender gap of approximately 26 log-point percent units in the baseline calibration. First, base productivity accounts for nearly half

[17] Appendix D provides more details.

of the GIG. Intuitively, it can be interpreted as a reduced-form combination of gender differences in labor-market productivity, wage-setting, and/or residual discrimination not otherwise captured by measured skills and choices. Second, Differences in the distribution of major-specific ability (means and dispersions) explain a large additional portion (around 8 percentage points). This channel operates through both *direct* effects (ability raises wages given major) and *indirect* effects (ability affects the set of majors that are attractive and feasible under fixed cutoffs). Because major assignment is capacity constrained and admission is governed by cutoffs, shifting the female ability distribution changes equilibrium selection into majors and therefore the wage distribution. Third, differences in hours profiles add another sizable portion (around 7 percentage points), which reduces the accumulation of experience in the human-capital technology. Interestingly, differing college major preferences and admission cutoffs help mitigate the GIG, nearly offsetting the impact of hours worked. First, restoring female major tastes changes where women apply and enroll; because women have much greater taste dispersion than men and relatively strong estimated ability in some fields, female-specific tastes reduce mismatch and reallocate women toward majors with higher expected lifetime value. Second, restoring female admission cutoffs reinforces the same mechanism: with capacity-constrained major assignment, male cutoffs ration some women out of majors in which they are relatively strong, whereas the baseline female admission cutoffs improves their access to those fields. Hence, preferences and admission rules mitigate the gender gap through equilibrium re-sorting across majors, partially offsetting the productivity and hours channels. Finally, factors such as female fertility, accumulated work experience, and income shocks tend to amplify the GIG. As noted by various authors, childbirth often results in women accumulating less work experience than men, linking fertility to experience.

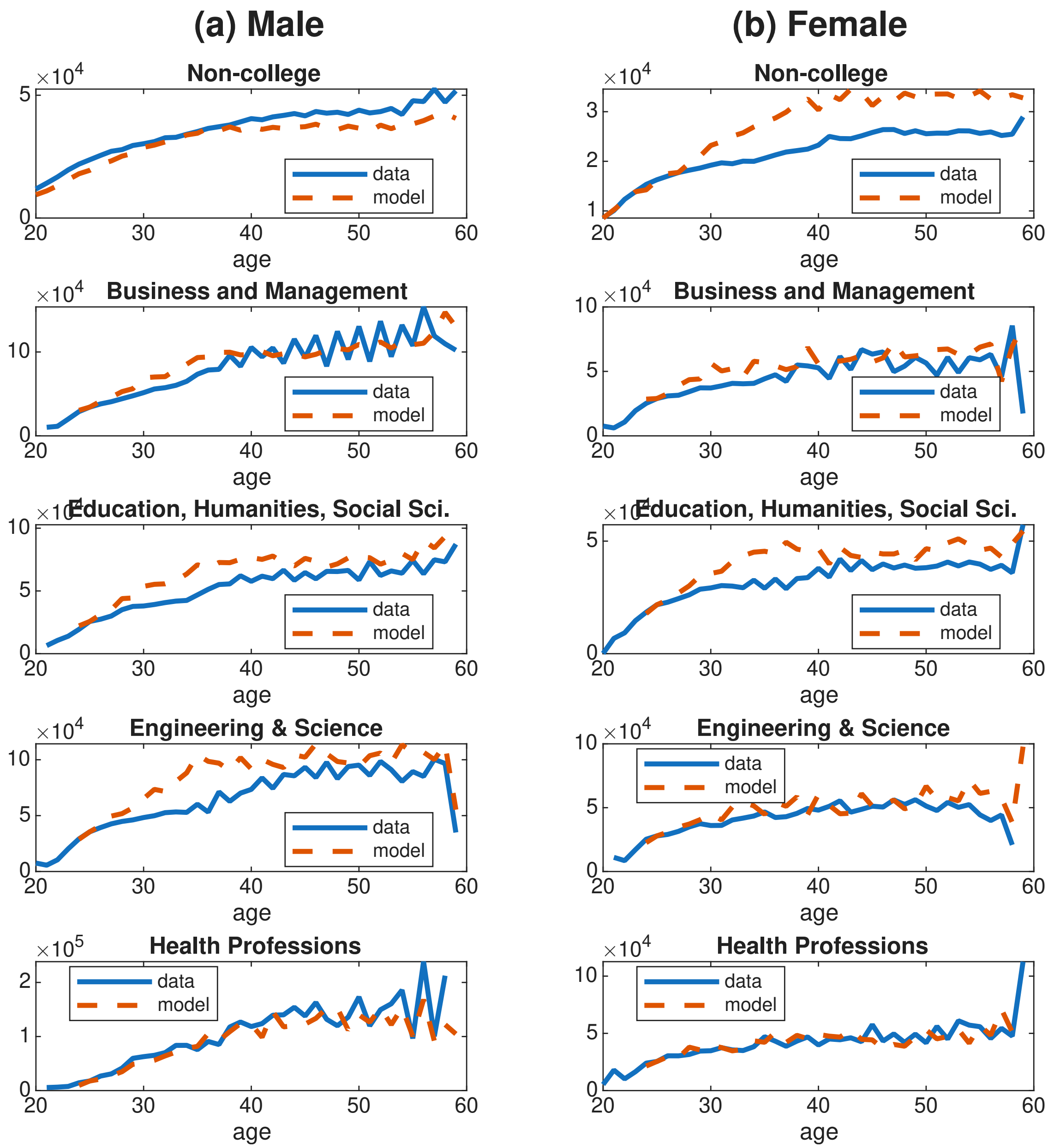

Figure 5: Life-cycle income profiles

Note: The figure compared the life-cycle income profiles (annual income in 2000 dollars) in the data vs. in the estimated model. The data is from the National Longitudinal Survey of Youth 1979 (NLSY79), 1979-2016.

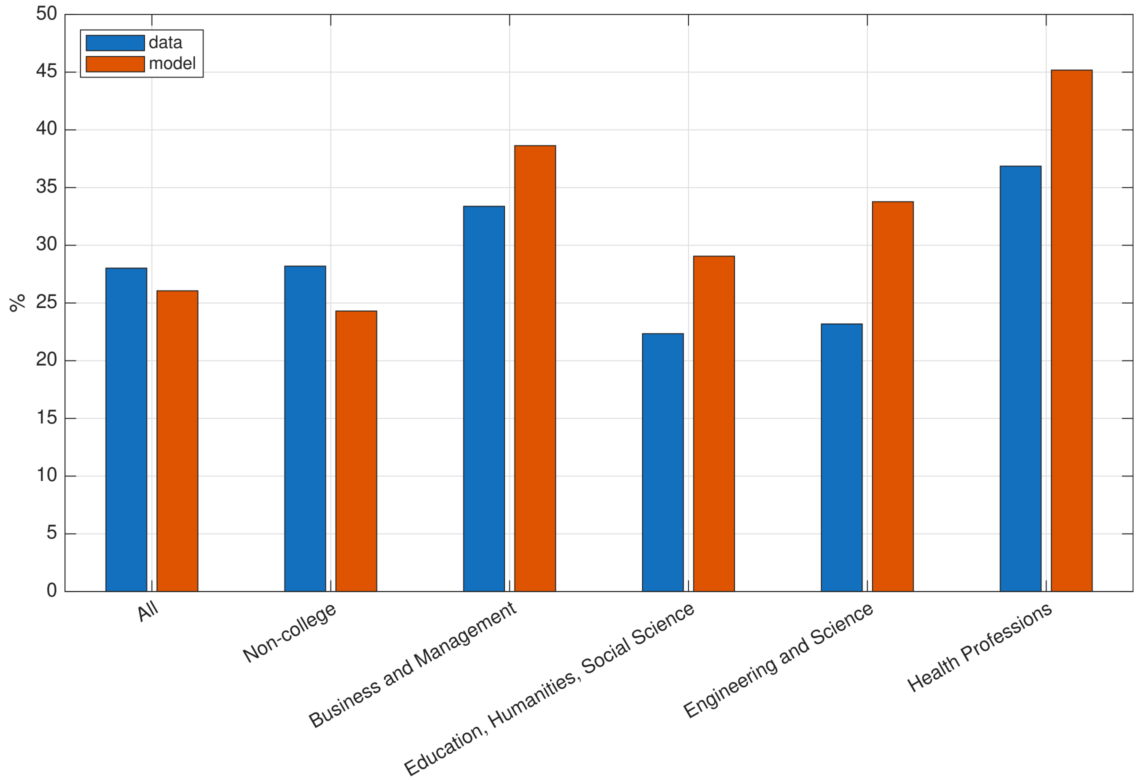

Figure 6: Gender gap: data vs. model

Note: The figure compared the gender gap (log difference in hourly wage) in the data vs. in the estimated model for each college major. The data is from the National Longitudinal Survey of Youth 1979 (NLSY79), 1979-2016.

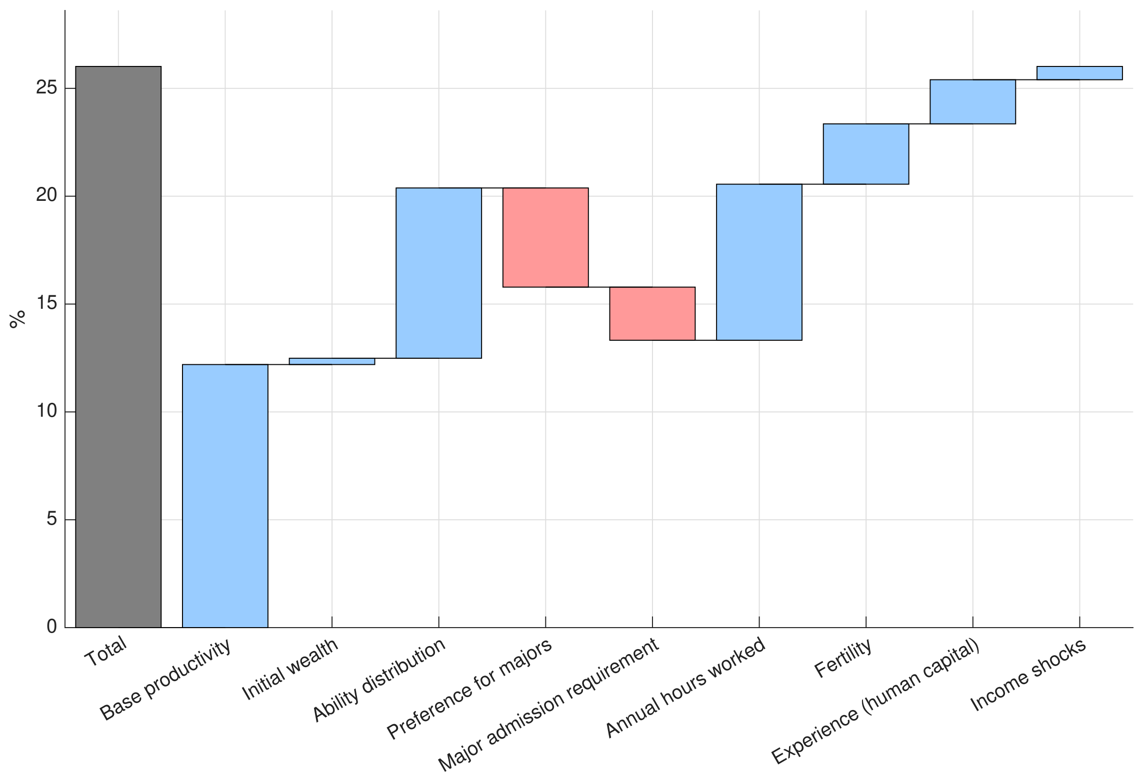

Figure 7: Decomposition of the Gender Gap

5 Policy Experiments and Counterfactual Analysis

In this section, we introduce several counterfactual policies. We will examine how different education policy regimes may affect the gender income gap, the gender welfare gap, the total welfare, the college enrollment, both the overall enrollment and each major-specific enrollment. We intend to draw lessons from the policy analysis. Appendix E contains additional results.

5.1 Uniform Tuition

The first policy regime we examine is uniform tuition, in which the same tuition rate is imposed across all majors despite varying per-student education costs. This policy scenario serves as a natural benchmark for two primary reasons. First, in practice, tuition rates across majors in the United States tend to be relatively uniform, with institutions rarely differentiating tuition substantially by field of study. Second, existing literature typically does not differentiate between college majors, effectively implementing a uniform tuition policy for all students. In this model, the government adjusts the progressive income tax schedule as it modifies the tuition rates to ensure that the budget remains balanced. Figure 8 presents the aggregate outcomes and Figure 9 presents the outcomes across graduates from different college majors. In the figures, recall that "unity" is the average initial wage of non-college workers for females with no child on an annual basis.

Several observations are in order. Figure 8 illustrates that total welfare is maximized at zero tuition. This is intuitive: with no tuition, individuals face no financial barriers to enrollment, making decisions solely based on preferences and expected future earnings. As tuition increases, overall welfare declines; individuals must evaluate whether college is "affordable." Even if initial wealth can cover tuition, the economic burden may deter some from enrolling, leading to a decrease in college attendance and, consequently, a decline in total welfare, given the positive college premium.[18] As tuition rises from zero to unity, the Gini coefficient increases.[19] Beyond this point, further increases in tuition have a relatively mild effect on the Gini coefficient.

A key insight from our analysis is *how gender disparities evolve with changes in tuition.* As tuition increases from zero to one, the gender income gap decreases while the gender welfare gap simultaneously widens. This suggests a trade-off between the income and welfare gaps at modest

[18] See Appendix E.2 for more details of the enrollment drop.

[19] In the baseline, tuition fees across majors range from 0.95 to 1.15 in model units.

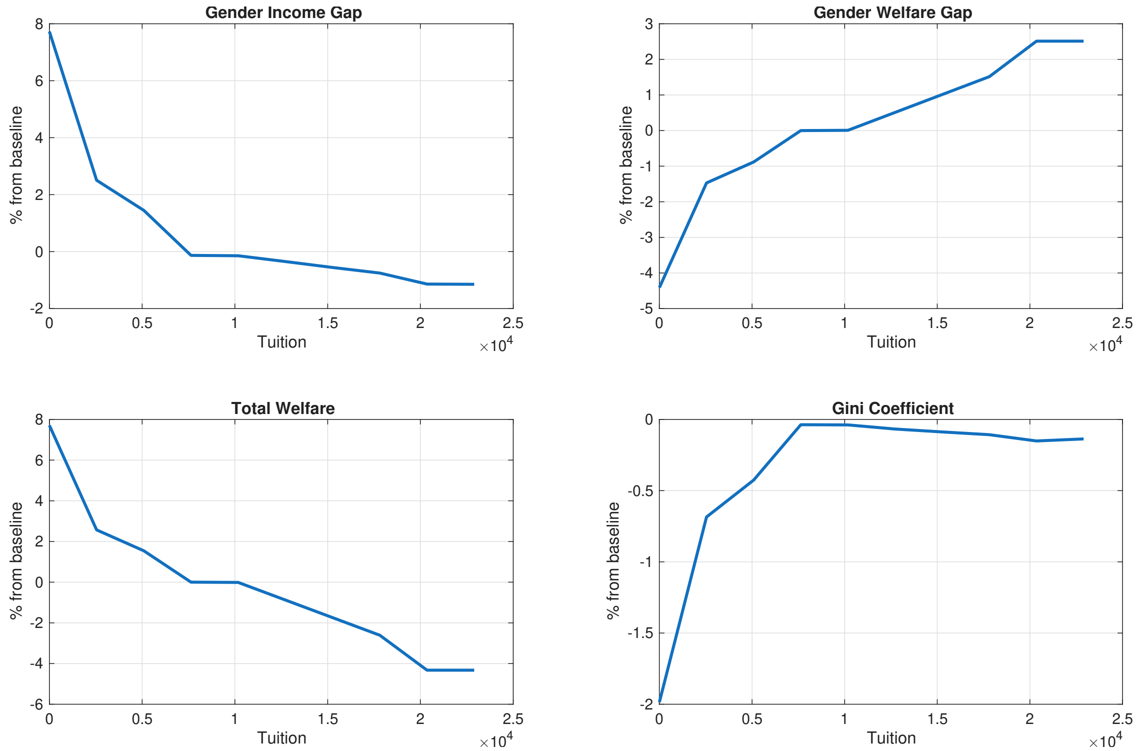

Figure 8: Uniform tuition policy

tuition levels close to the status quo.

Here are some intuitions of why the gender income gap decreases yet the gender welfare gap increases with the college tuition. First, the gender income gap among college graduates is larger than that among non-graduates. Consequently, when both men and women opt out of college, the income gap narrows—this is the "extensive margin effect." Second, the "intensive margin effect" comes into play: individuals experience taste shocks for each major, and estimates show that women have greater variance in these shocks than men. As tuition rises, men gain a comparative advantage in balancing the choice between a "more preferred" major and a "more profitable" one.

5.2 Randomization of College Majors

Countries vary significantly in their university admission processes. Some allow students to enroll before selecting a major, while others have major-specific admissions, accepting students into designated majors at particular institutions (Bordon and Fu (2015)). In the latter case, both applicants and universities depend on "signals" from high school to guide their decisions. This

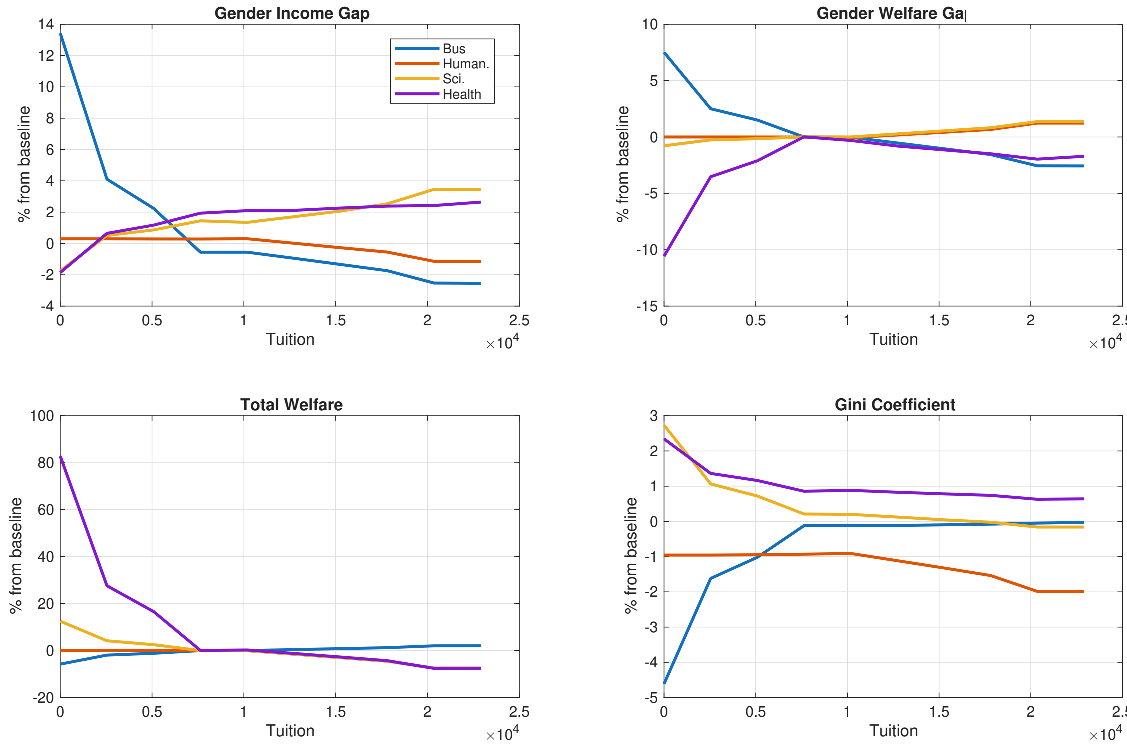

Figure 9: Uniform tuition policy (by major)

raises an important question: what occurs when these signals are incomplete? For instance, some college majors may not be adequately represented in the high school curriculum. Moreover, how can high school students effectively differentiate between fields such as biology, chemistry, biochemistry, and pharmacy?

Here we consider an extreme scenario, assuming that all high school signals are uninformative. In this case, college majors are effectively allocated to students at random.[20] We examine this extreme situation to underscore the importance of college major choice.

Figure 10 illustrates that the impact on the Gini coefficient, a measure of income inequality, is relatively small. However, total welfare decreases by more than 30%. Notably, the gender income gap has minimal changes, and the gender welfare gap increases by about 30% compared to the baseline scenario. Recall that females tend to experience *greater* taste shocks related to their college majors. Therefore, random allocation of college majors would likely discourage female students from enrolling in college more than their male counterparts, thereby *exacerbating* the gender welfare gap.

[20]Specifically, we fix the same pool of college-attending students as in the baseline, then randomize their majors.

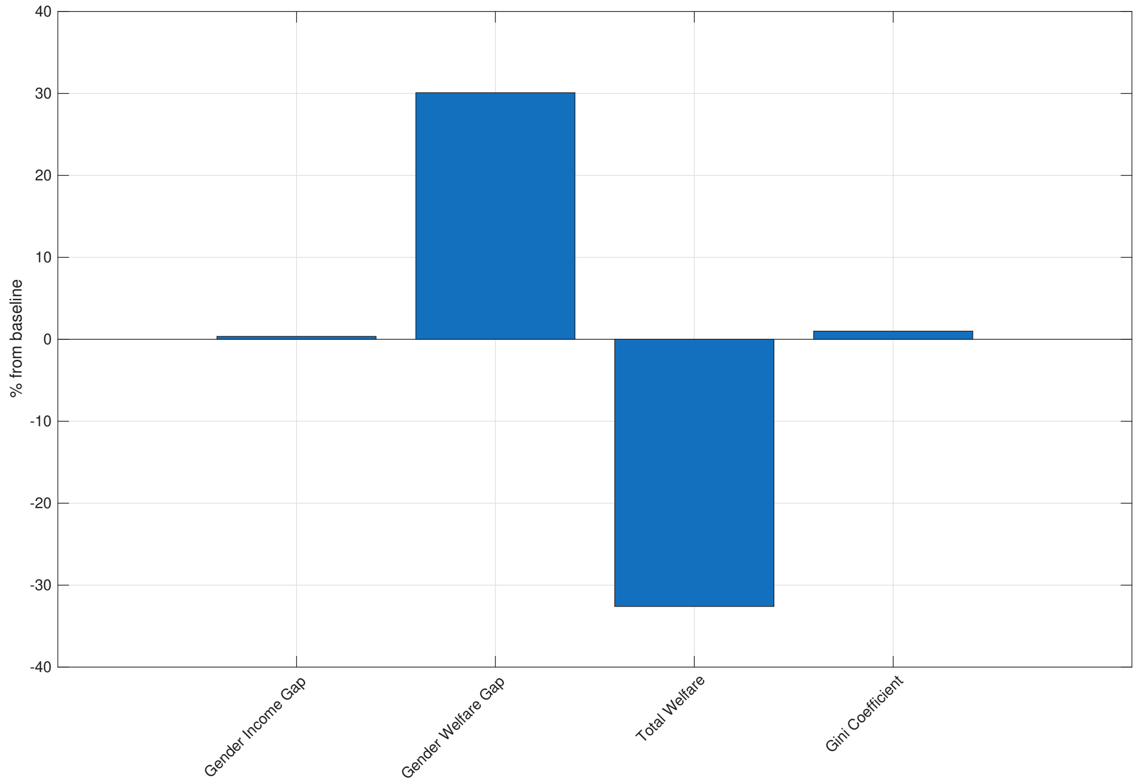

Figure 10: Randomization of college majors

5.3 Gender-based Affirmative Action

There is a substantial body of literature on racial affirmative action (Akhtari, Bau and Laliberté (2024)). In this section, we apply the concept of affirmative action to gender major choice. Specifically, we require each college major to enroll an equal ratio of gender types as in the population to ensure equal representation, and we examine the potential impacts on economic outcomes.

Figure 11 provides a visual representation of these effects. Gender-based affirmative action has minimal impact on gender income gap and the Gini coefficient. However, its *distributional* effects are significant: the gender welfare gap increases by almost 10%, while the total welfare decreases by nearly 15%. The underlying intuition remains the same as before: females derive more utility from their college majors than males. Thus, the welfare loss of males are less than females under the gender-based affirmative action, and hence the gender welfare gap increases.

Figure 12 offers further insights. Gender-based affirmative action distorts the *composition* of female students. Notably, compared to the baseline scenario, the college enrollment rate for females

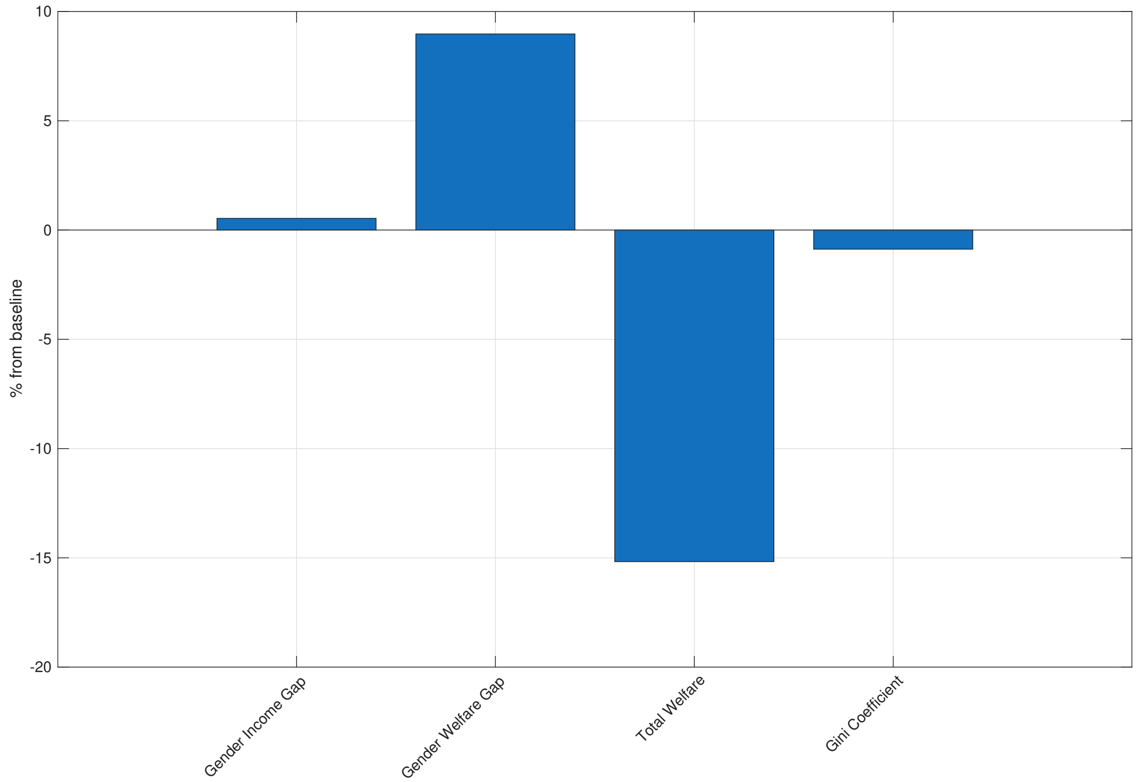

Figure 11: Gender-based affirmative action

drops by over 2%, primarily affecting those in the humanities and health majors. This is partially offset by an increase in enrollment in business and science fields.

An important lesson from this exercise is that the gender income gap and the gender welfare gap can move in opposite directions. Policymakers aiming to promote gender equality should be mindful of this subtlety.

5.4 Tradeoff Between Income and Welfare Gaps under different tuition policies

This paper builds on the insightful observation made by Altonji and Zimmerman (2019) that the per-student cost varies significantly across different college majors. Thus, uniform tuition is not the only viable option. For instance, the government could subsidize various college majors by the same percentage. In this scenario, while students would pay different amounts of tuition depending on their major, the ratio of tuition paid to the true cost of education would be equalized across majors. We refer to this as the ”uniform subsidy rate” regime. Alternatively, one could envision a system where every college student pays the full cost associated with their major, and

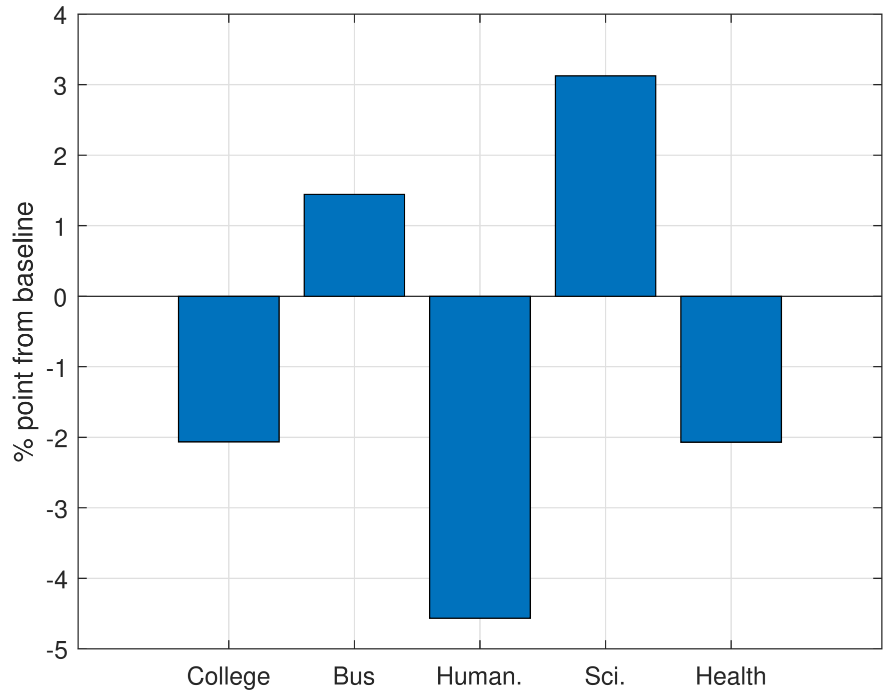

Figure 12: Changes in enrollment for females due to affirmative action

the government reimburses an equal amount to all students. This concept is akin to the idea of education coupons (Friedman (1962)). We label this regime as the ”uniform subsidy per capita,” or ”lump sum transfer” regime. In a way, the lump sum transfer regime is like giving every college student *an education voucher* and allow them to make their own choices.

Our analysis of the uniform tuition subsidy shows a negative correlation between the gender income gap and the gender welfare gap at modest tuition levels (see Figure 8). It's crucial to note that we can only directly measure the gender income gap; the gender welfare gap must be evaluated through a structural model, like the one presented in this paper. This presents a significant challenge for policymakers, as efforts to reduce the gender income gap may unintentionally exacerbate the gender welfare gap.

We investigate whether the relationship between the gender income gap and the gender welfare gap holds under different policy regimes. Beyond the uniform tuition regime, we will also evaluate uniform subsidy ratios and lump sum transfer regimes. Figure 13 summarizes this comparison. Generally, for a given gender income gap, the uniform tuition regime results in a lower gender welfare gap, except when tuition for the lowest-cost major approaches its actual cost. In this context, the uniform tuition regime outperforms the others. Moreover, the relationship between the gender income gap and the gender welfare gap is consistently negative across all regimes. Therefore, we conclude that the empirically plausible relationship between the gender income gap and the gender welfare gap can be *negative*, suggesting that closing the gender income gap may inadvertently widen the welfare gap between genders.

5.5 Tradeoff Between Efficiency and Equity

Policymakers confront *various trade-offs*, including the gender income gap versus the welfare gap studies in the previous section. In this section, we examine the trade-off between efficiency and equality, treating them as distinct "social goods" with different policy regimes acting as varying "production technologies." Our goal is to identify if any single regime can achieve a superior production frontier for social goods. We measure efficiency as aggregated expected lifetime utility (or "welfare") and assess equality using the Gini ratio (1 - Gini).[21] We also analyze how efficiency and equality would trade-off with the gender gap in income and utility.

Figure 14 provides a graphical illustration of the comparison, with the baseline case marked as a star in the figure. Panel (a) reveals that at relatively lower levels of efficiency, the uniform tuition regime achieves a higher level of income equality. Conversely, at higher levels of efficiency, the lump sum transfer regime offers greater income equality. Therefore, no single regime can completely dominate the others.

However, Panel (b) indicates that at relatively lower levels of efficiency, different policy regimes provide similar levels of welfare equality. At higher levels of efficiency, the lump sum transfer regime achieves a higher level of income equality. Therefore, if one prioritizes welfare equality, the lump sum transfer regime may be the preferred option.

5.6 The Role of College Majors

This paper develops a framework incorporating college major choices and gender differences, revealing the perhaps unexpected result that the gender income gap is negatively related to the gender welfare gap. A key question arises regarding how college major choices and gender differences

[21] Fehr and Charness (2025) reviews the literature, concluding that individuals care about both their own payoffs and those of others, reflecting social preferences. Our metrics align with studies such as Hanushek, Leung and Yilmaz (2003) and Guo and Leung (2021).

shape the relationship between gender income disparities, welfare gaps, and overall economic welfare. To explore this, we first analyze a simplified version of our baseline model, where individuals face a binary choice: to attend college or not, omitting specific major choices.

Figure 15 illustrates how the relationships between gender income gaps, gender welfare gaps, and total welfare differ under uniform tuition policy regimes. The negative relationship between the gender income gap and the gender welfare gap largely disappears once we abstract from major choice and collapse college into a single alternative (college vs. non-college). In fact, under different tuition policies, the two gap measures are now largely consistent with each other. Intuitively, when there is no major choice, tuition policy no longer changes the relative price of different types of human capital. The main behavioral margin is the extensive margin of college attendance, so the policy operates primarily through a extensive margin channel rather than through re-sorting across fields. In that environment, higher earnings for women from increased college participation are much more tightly linked to higher utility: there is no major mismatch margin through which earnings gains are offset by losses in non-pecuniary surplus. Consequently, policies that reduce the income gap tend also to reduce the welfare gap, and the tradeoff observed in Figure 13 is no longer present in Figure 15.

This exercise highlights the important fact that accounting for major heterogeneity is not a quantitative detail but a qualitative one. The fact that the income–welfare tradeoff vanishes in the no-major benchmark underscores that the structure of within-college allocation—who studies what, and under what subsidy schedule—is essential for understanding the distributional consequences of higher-education finance reforms by gender.

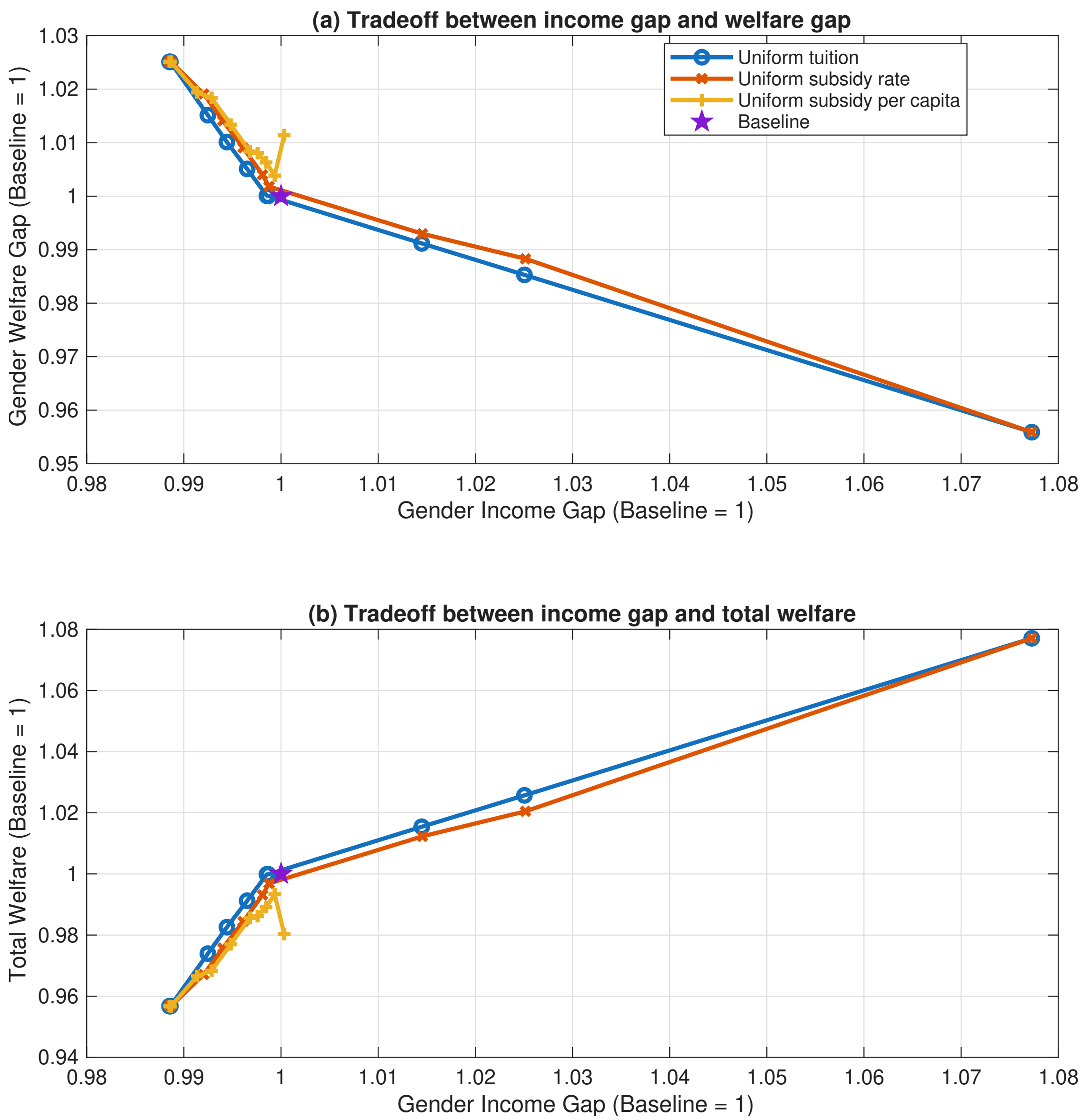

Figure 13: Tradeoff Between Income and Welfare Gaps

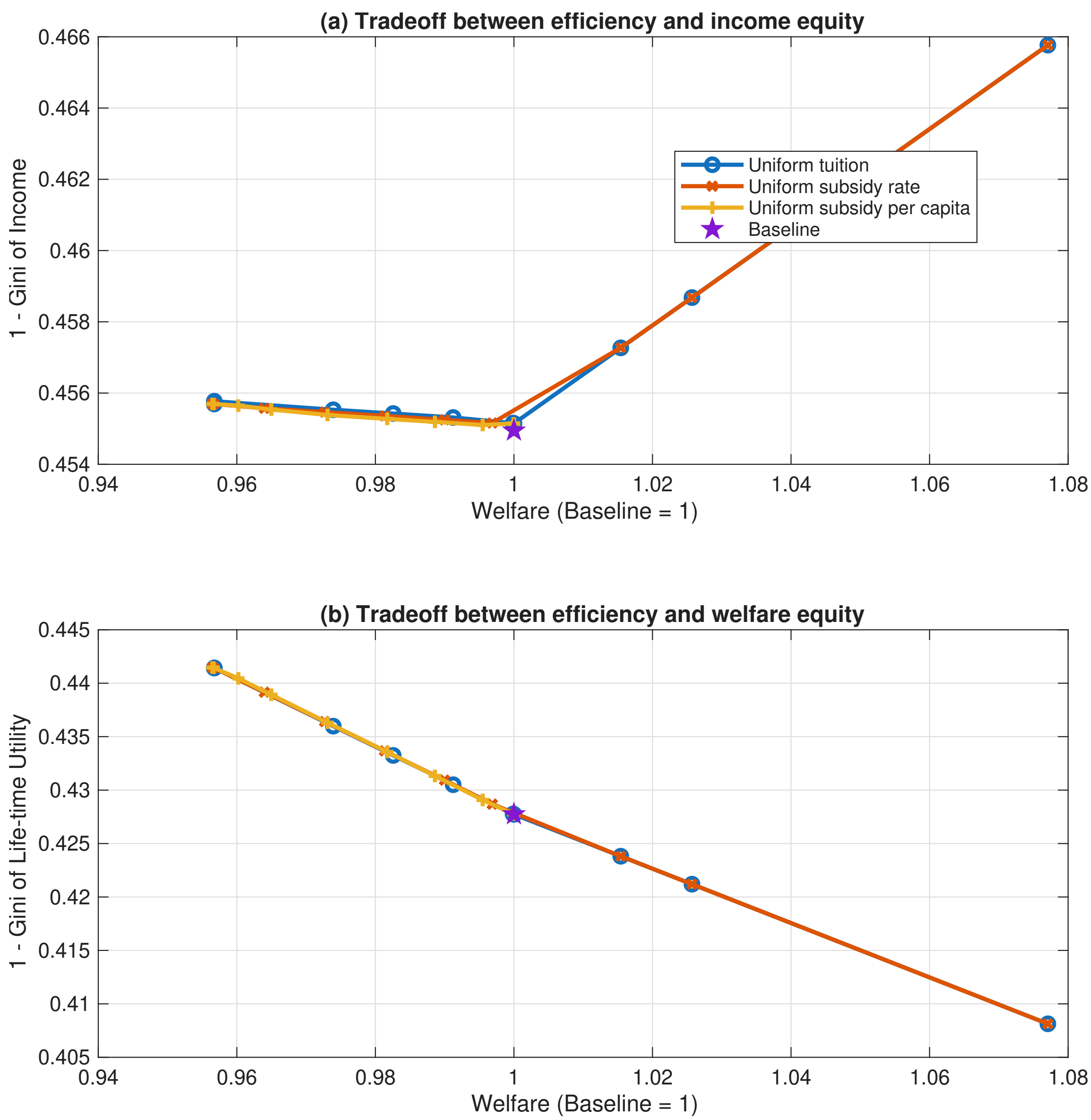

Figure 14: Tradeoff between efficiency and equity

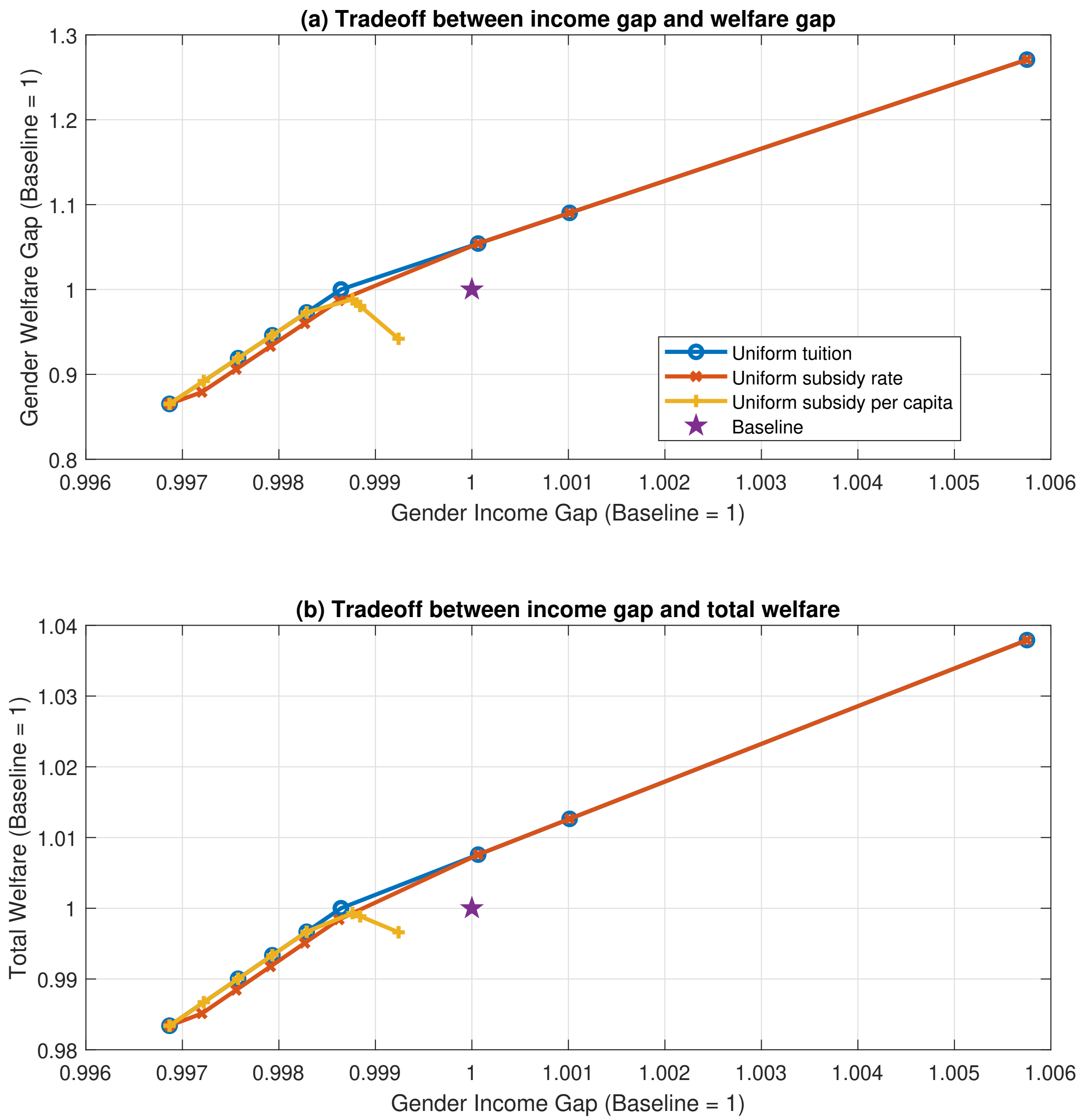

Figure 15: Tradeoffs with no major choice

6 Conclusion

This paper develops a quantitative life-cycle model to investigate how educational policies influence gender inequality by explicitly considering endogenous college major choices. We highlight the importance of distinguishing between two dimensions of gender disparities: the income gap and the welfare gap. Using the calibrated model, we quantitatively evaluate a range of educational policy scenarios, including uniform tuition, gender-based affirmative action, random major assignment, etc.

Our findings reveal significant and nuanced trade-offs between the gender income gap and the gender welfare gap. For instance, universal free tuition significantly improves welfare by removing financial constraints, but it simultaneously exacerbates gender income disparities. In contrast, both gender-based affirmative action aimed at achieving equal representation across majors, and randomly assigning college majors, have little impact on the gender income gap, yet substantially worsen gender welfare disparities, reflecting tensions between economic returns and individual preferences.

Further, our analysis underscores the critical role of explicitly modeling heterogeneity across college majors and gender. Simplified models abstracting from detailed major choices fail to capture crucial aspects of the relationship between gender gaps and overall welfare, potentially leading policymakers astray. In particular, neglecting detailed college-major choices systematically obscures the critical trade-off netween the income gap and the welfare gap.

The results highlight the complexity of gender disparities in education and labor markets, emphasizing that policies aimed solely at narrowing income gaps may inadvertently compromise welfare outcomes, and vice versa. Thus, achieving gender equality through education policy requires carefully balancing economic returns and individual preferences, acknowledging nuanced interactions between policy design, individual choices, and labor market outcomes.

References

Akhtari, Mitra, Natalie Bau, and Jean Laliberté. 2024. "Affirmative Action and Precollege Human Capital." *American Economic Journal: Applied Economics*, 16(1): 1–32.

Altonji, Joseph, and Seth Zimmerman. 2019. "The Costs of and Net Returns to College Major." *Productivity in Higher Education*, , ed. Caroline M. Hoxby, Kevin Strange and Kevin Stange, 133–176. Chicago:University of Chicago Press.

Altonji, Joseph G., Erica Blom, and Costas Meghir. 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers." *Annual Review of Economics*, 4(Volume 4, 2012): 185–223.

Altonji, Joseph, Lisa Kahn, and Jamin Speer. 2016. "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success." *Journal of Labor Economics*, 34(S1): 361–401.

Altonji, Joseph, Peter Arcidiacono, and Arnaud Maurel. 2016. "The Analysis of Field Choice in College and Graduate School." *Handbook of the Economics of Education, Volume 5*, , ed. Eric Hanushek, Stephen Machin and Ludger Woessmann, 305–396. New York:Elsevier.

Andrews, Rodney, Scott Imberman, Michael Lovenheim, and Kevin Stange. 2024. "The Returns to College Major Choice: Average and Distributional Effects, Career Trajectories." NBER Working Paper No. 30331.

Angrist, Joshua, David Autor, and Amanda Pallais. 2023. "Marginal Effects of Merit Aid for Low-Income Students." *Quarterly Journal of Economics*, 137(2): 1039–1090.

Arcidiacono, Peter. 2004. "Ability sorting and the returns to college major." *Journal of Econometrics*, 121(1-2): 343–375.

Arcidiacono, Peter, V. Joseph Hotz, and Songman Kang. 2012. "Modeling college major choices using elicited measures of expectations and counterfactuals." *Journal of Econometrics*, 166(1): 3–16.

Attanasio, Orazio P., and Guglielmo Weber. 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy*, 103(6): 1121–1157.

Bang, Minji. 2021. "Job Flexibility and Household Labor Supply: Understanding Gender Gaps and the Child Wage Penalty." Working Paper.

Barth, Erling, Sari Pekkala Kerr, and Claudia Olivetti. 2021. "The dynamics of gender earnings differentials: Evidence from establishment data." *European Economic Review*, 134 Article 103713.

Bick, Alexander, Adam Blandin, and Richard Rogerson. 2025. "After 40 Years, How Representative Are Labor Market Outcomes in the NLSY79?" *Federal Reserve Bank of St. Louis Review*, 107(2): 1–50.

Black, Sandra, Jeffrey Denning, Lisa Dettling, Sarena Goodman, and Lesley Turner. 2023. "Taking It to the Limit: Effects of Increased Student Loan Availability on Attainment, Earnings, and Financial Well-Being." *American Economic Review*, 113(12): 3357–3400.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature*, 55(3): 789–865.

Bordon, Paola, and Chao Fu. 2015. "College-Major Choice to College-Then-Major Choice." *Review of Economic Studies*, 82(4): 1247–1288.

Bulman, George, Robert Fairlie, Sarena Goodman, and Adam Isen. 2021. "Parental Resources and College Attendance: Evidence from Lottery Wins." *American Economic Review*, 111(4): 1201–1240.

Caucutt, Elizabeth, and Lance Lochner. 2020. "Early and Late Human Capital Investments, Borrowing Constraints, and the Family." *Journal of Political Economy*, 128(3): 1065–1147.

Choi, Woosuk, Josh Kinsler, Alexis Orellana, and Ronni Pavan. 2023. "College Majors and Earnings Growth." Working Paper.

Darolia, Rajeev, and Dubravka Ritter. 2020. "Strategic Default Among Private Student Loan Debtors: Evidence from Bankruptcy Reform." *Education Finance and Policy*, 15(3): 487–517.

Dynarski, Susan, Lindsay Page, and Judith Judith Scott-Clayton. 2022. "College Costs, Financial Aid, and Student Decisions." *NBER Working Paper No. 30275.*

Erosa, Andres, Luisa Fuster, and Diego Restuccia. 2016. "A quantitative theory of the gender gap in wages." *European Economic Review*, 85: 165–187.

Erosa, Andres, Luisa Fuster, Gueorgui Kambourov, and Richard Rogerson. 2022. "Hours, Occupations, and Gender Differences in Labor Market Outcomes." *American Economic Journal: Macroeconomics*, 14: 543–590.

Fehr, Ernst, and Gary Charness. 2025. "Social Preferences: Fundamental Characteristics and Economic Consequences." *Journal of Economic Literature*, 63(2): 440–514.

Friedman, Milton. 1955. "The Role of Government in Education." *Economics and the Public Interest*, , ed. Robert Solo, 305–396. New Jeresy:Rutgers University Press.

Friedman, Milton. 1962. *Capitalism and Freedom.* Chicago: University of Chicago Press.

Gale, D., and L. S. Shapley. 1962. "College Admissions and the Stability of Marriage." *The American Mathematical Monthly*, 69(1): 9–15.

Gemici, Ahu, and Matthew Wiswall. 2014. "Evolution Of Gender Differences In Post-Secondary Human Capital Investments: College Majors." *International Economic Review*, 55(1): 23–56.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review*, 104(4): 1091–1119.

Goldin, Claudia, Sari Pekkala Kerr, and Claudia Olivetti. 2024. "The parental pay gap over the life cycle: Children, jobs, and labor supply." *Journal of Economic Dynamics and Control, Article 104963.*

Goldin, Claudia, Sari Pekkala Kerr, Claudia Olivetti, and Erling Barth. 2017. "The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census." *American Economic Review*, 107(5): 110–114.

Gong, Yifan, Lance Lochner, Ralph Stinebrickner, and Todd Stinebrickner. 2018. "The consumption value of college." Working paper.

Gottlieb, Joshua, Neale Mahoney, Kevin Rinz, and Victoria Udalova. 2025. "The Rise of Healthcare Jobs." NBER Working Paper No. 33583.

Guner, Nezih, Martin Lopez-Daneri, and Gustavo Ventura. 2016. "Heterogeneity and

Government revenues: Higher taxes at the top?" *Journal of Monetary Economics*, 80: 69–85.

Guo, Naijia, and Charles Ka Yui Leung. 2021. "Do elite colleges matter? The impact on entrepreneurship decisions and career dynamics." *Quantitative Economics*, 12: 1347–1397.

Hanushek, Eric, Charles Ka Yui Leung, and Kuzey Yilmaz. 2003. "Redistribution through education and other transfer mechanisms." *Journal of Monetary Economics*, 50(8): 1719–1750.

Jacob, Brian, Brian McCall, and Kevin Stange. 2018. "College as Country Club: Do Colleges Cater to Students' Preferences for Consumption?" *Journal of Labor Economics*, 36(2): 309–348.

Keane, Michael P, and Kenneth I Wolpin. 1997. "The career decisions of young men." *Journal of Political Economy*, 105(3): 473–522.

Keane, Michael, Petra Todd, and Kenneth Wolpin. 2011. "The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications." *Handbook of Labor Economics, volume 4*, , ed. O. Ashenfelter and D. Card, 331–461. New York:Elsevier.

Lee, Sang Yoon, and Ananth Seshadri. 2019. "On the intergenerational transmission of economic status." *Journal of Political Economy*, 127(2): 855–921.

Leighton, Margaret, and Jamin Speer. 2023. "Rich Grad, Poor Grad: Family Background and College Major Choice." IZA Working Paper No. 16099.

Matsuda, Kazushige, and Karol Mazur. 2022. "College education and income contingent loans in equilibrium." *Journal of Monetary Economics*, 132(C): 100–117.

OECD. 2023. "Reporting Gender Pay Gaps in OECD Countries." *https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/06/reporting-gender-pay-gaps-in-oecd-countries_82110f1a/ea13aa68-en.pdf*.

Patnaik, Arpita, Joanna Venator, Matthew Wiswall, and Basit Zafar. 2022. "The role of heterogeneous risk preferences, discount rates, and earnings expectations in college major choice." *Journal of Econometrics*, 231(1): 98–122.

Sjoquist, David, and John Winters. 2015. "State Merit Aid Programs and College Major: A Focus on STEM." *Journal of Labor Economics*, 33(4): 973–1006.

Stinebrickner, Ralph, and Todd Stinebrickner. 2014. "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout." *Review of Economic Studies*, 81(1): 426–472.