ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

On Expanding Public Funding of Selective Colleges

Authors	Lutz Hendricks, Tatyana Koreshkova, and Oksana Leukhina
Working Paper Number	2025-011A
Creation Date	May 2025
Citable Link	https://doi.org/10.20955/wp.2025.011
Suggested Citation	Hendricks, L., Koreshkova, T., Leukhina, O., 2025; On Expanding Public Funding of Selective Colleges, Federal Reserve Bank of St. Louis Working Paper 2025-011. URL https://doi.org/10.20955/wp.2025.011

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

On Expanding Public Funding of Selective Colleges

Lutz Hendricks*

Tatyana Koreshkova†

Oksana Leukhina‡

March 20, 2025

Click here for the latest version

Abstract

As college attainment expanded in the U.S., the fraction of public funds allocated to selective colleges and universities declined. Does this make sense from an efficiency standpoint, given that the majority of college entrants face the highest financial returns at selective colleges? Should the states instead be expanding access to high quality colleges? In this paper, we examine reallocating public funds from low quality to high quality colleges in a spending-neutral way. We find that this policy leads to a decline in aggregate earnings and intergenerational income mobility. Public spending on lower quality schools is well-justified.

1 Introduction

Most college students in the U.S. attend institutions that are publicly operated and, at least in part, publicly funded. These institutions differ greatly in "quality." Some are four-year institutions that grant bachelor's degrees; others are two-year colleges that do not. Some

*Department of Economics, University of North Carolina, Chapel Hill, NC, USA. Email: hendricksl@protonmail.com.

†Department of Economics, Concordia University, Montreal, Canada. Email: tatyana.koreshkova@concordia.ca.

‡Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166, USA. Email: oksana.m.leukhina@gmail.com.

are highly selective "flagship" universities; others are less selective and cater to less prepared students.

Our broad research question is: How should public funds be allocated across institutions of different qualities? As educational attainment has expanded in the past, public funds have shifted towards less selective institutions and, especially, towards two-year colleges (Bastedo and Jaquette, 2011). Was this reallocation of funds optimal? To our knowledge, these questions have not been addressed in previous research.

Consider a government that plans to spend a given amount of funds on expanding the number of college seats. Should these funds be directed towards the highly selective "flagship" colleges or towards the less selective colleges?

The answer is not obvious ex-ante. The government faces a quantity-quality trade-off. On one hand, each additional seat added to a high-quality school is more likely to produce a college graduate and offers a larger earnings gain upon graduation (Hoekstra, 2020; Lovenheim and Smith, 2023; Kinsler and Pavan, 2011). On the other hand, creating additional seats in higher-quality colleges is more expensive – instructional expenditures per student increase sharply with college quality (Hoxby, 2009, Figure 2). Thus, a given amount of public funds will create fewer college seats in higher-quality colleges.

Improving the allocation of funds across colleges has the potential to generate large welfare gains. High quality colleges face substantial excess demand and therefore ration out a large fraction of applicants. Additional seats in these colleges would likely get filled, thereby benefitting the pulled-in students and potentially raising aggregate human capital and earnings. This sentiment is, in fact, reflected in funding rules adopted by many states. Performance-based funding models, for example, tie public funds to graduation rates, major compositions, or post-graduation outcomes.

We consider four types of colleges. In Section Section 3.1.1, we explain how we categorize all colleges and universities in the U.S. into four quality groups that we model ($q \in \{1, 2, 3, 4\}$). We show that higher quality colleges have higher graduation rates and their graduates earn more. Higher quality colleges have higher graduation rates and their graduates earn more: compared to Quality 2 colleges, Quality 3- (Quality 4-) graduates earn about 10% (30%) more over the lifetime.

The paper's objective is to examine the implications of reallocation public funds from the least selective four-year colleges ($q2$) to the most selective four-year colleges ($q4$). To this end, we develop a model of college quality choice and human capital accumulation in college. The model is needed so as to understand how different kinds of students resort across colleges in response to changes in public finance.

Our model features rich heterogeneity among students. Students enter the model at high school graduation and differ in terms of family income, learning ability, test scores, initial level of human capital and their preferences for each college group. Students choose between working as a high school graduate and enrolling in one of the colleges they can get admitted to. In light of existing evidence on prevalence of information frictions, we also assume that students face a degree of uncertainty about college quality. While in college, students accumulate human capital and face graduation uncertainty. Both matter for their earnings upon labor market entry.

Colleges differ in human capital accumulation technology and dropout risk, both of which interact with student ability. Human capital accumulation technology captures learning and network opportunities as well as instructional quality. Dropout risk captures advising quality, instruction quality and stringency of graduation requirements. Colleges also differ in their financial costs. Higher quality colleges tend to charge higher tuition. Parental transfers increase with college quality, but much more so for the high family income students.

Four-year colleges have fixed capacities, and admissions are determined as an equilibrium outcome in the college market. We do not explicitly model the optimization problem solved by colleges, instead assuming they maximize some measure of their students' achievement, subject to operating at full capacity. We postulate this measure of achievement is a combination of both, the test score and initial human capital, calibrating their relative importance. As students make their college selection in order of their achievement rank, colleges get gradually filled – college-specific admission cutoffs thus determined by the rank of the student taking the last available spot. It follows that lower ranking students may be rationed out of their preferred college by admissions.

We discipline the model by targeting detailed characteristics of the NLSY97 cohort as well as two elasticities based on quasi-experimental evidence. We augment the publicly available NLSY97 data file with Geocode data and official college transcripts. All of the financial variables associated with college are taken directly from the data and therefore reflect all the relevant empirical features such as prevalence of merit-based scholarships, financial aid and parental willingness to subsidize colleges of different quality. Our intention is to ensure that the model accurately matches consumption levels for different students at different colleges.

The earnings targets help us identify college human capital accumulation technologies as well as the complementarity between student ability and college quality. Dropout and graduation rates help identify dropout risk. Together with college-related financial variables and human capital technologies, these determine financial returns to different colleges for different type of students.

The calibrated model successfully matches the targets. We find that higher quality colleges offer a more productive human capital accumulation technology which also features a stronger degree of complementarity with student ability. This means that access to high quality learning technology is particularly important for high ability students, and they are less likely to be deterred from high quality schools by financial constraints or preference draws. Nonetheless, most students face higher returns in higher quality schools and would choose Quality 3 or Quality 4 schools based on financial returns alone.

We find that reallocating a fixed amount of public funds from low quality to high quality four-year colleges leads to a decline of both aggregate earnings and upward income mobility. The intuition is as follows. For every seat created in high quality colleges, a little over two seats are lost in low quality colleges – because per-student expenditures are substantially lower at these schools. Fewer students enter college because of the overall loss of capacity. The number of college graduates, however, remains roughly unchanged because graduation rates increase with greater access to high quality schools. While the graduates do earn more, consistent with our initial intuition, the loss of access to lower quality seats implies fewer dropouts and therefore lower aggregate earnings.

1.1 Related Literature

To our knowledge, this is the first paper that studies the allocation of funds to colleges of different qualities. As pointed out by Lovenheim and Smith (2023), the literature has studied the gross returns to college quality (the implied earnings gains) in great detail. Far less is known about the net returns to public investment.

1.1.1 College Quality and Earnings

An empirical literature studies the implications of attending "better" colleges for earnings later in life (see Hoekstra 2020 and Lovenheim and Smith 2023 for surveys). We draw on one of these studies (Hoekstra, 2009) for our calibration. Papers with strong identification often rely on discontinuities in admissions rules. Notably, Texas's "Top 10 Percent Rule" automatically admits the top ten percent of graduates from qualifying Texas high schools to the state's top public universities (Black et al., 2023b). Similarly, California's "Eligibility in the Local Context" grants admission to some selective colleges for students in the top four percent of their high school class (Bleemer, 2024). Admission discontinuities allow researchers to identify the causal effects of attending more selective colleges based on plausible and clearly stated assumptions. Our quantitative model is broadly consistent with these findings.

2 Model

2.1 Model Overview

This paper aims to quantify how improving college access for lower income students affects intergenerational earnings mobility and aggregate earnings. For this purpose, we develop a model of college choice with the following key features:

1. The model follows a single cohort of high school graduates through college and work into retirement.
2. High school graduates differ in their learning ability and human capital, which affect the financial returns to college. They also differ in terms of parental background which determines their ability to pay for college.
3. Colleges are places of learning that differ in terms of "quality" q. Better colleges produce more human capital, at least for high ability students. They also charge higher tuition.
4. Colleges are capacity constrained. For each additional low income student that enrolls, one high income student is displaced.
5. Students, especially from low income backgrounds, face various frictions when selecting colleges:
   (a) Financial constraints: Students have limited resources to pay for potentially expensive selective colleges.
   (b) Admissions: High quality colleges practice selective admissions. Low income students tend to present with inferior resumes (in the model: lower initial human capital) and are therefore admitted at low rates.
   (c) Information: Students are imperfectly informed about the financial returns to colleges of different quality levels.[1]
   (d) Preferences: Students have idiosyncratic preferences for specific colleges. Evidence suggests that many students enroll in colleges that are either close to home or that are attended by friends or peers (Armstrong, 2013; Dillon and Smith, 2017).

   Jointly, these frictions generate undermatch, especially among low income students. The undermatch driven by admissions constraints form the pool of students who may be induced to enroll in better colleges when new spots are created there.

The timing of events is as follows:

[1] "Young people—particularly those from lower-income, immigrant, and/or non-college educated families—may lack good information about the costs and benefits of enrollment, the process of preparing for, applying to, and selecting a college" (Dynarski et al., 2022a, p. 3).

1. Students draw endowments (ability, parental background, etc.). The endowments imply admissions scores z.
2. College q admits all students with admissions scores above the cutoff value $\bar{z}_q$. Colleges have limited capacities and set the cutoff values so as to fill all available seats.
3. Students choose a college from the set they are admitted to; or they start to work as high school graduates. At this stage, students imperfectly observe the quality of admitting colleges.
4. In each college period, students accumulate human capital. The rate of learning depends on student ability and on college quality. Students also consume and borrow.
5. At the end of each college period, students may drop out or graduate, in which case they become workers.
6. After completing their education, workers solve a simple permanent income consumption-saving problem. Worker earnings are determined by the human capital accumulated in college and by degree attainment (a sheepskin effect).

The following sections describe these model stages in detail. We discuss our modeling choices in section Section 2.8.

2.2 Student Endowments

High school graduates enter the model at age $t = 1$ (physical age 19). They draw a vector of endowments that consists of learning ability a (standard Normal marginal), parental income percentile p, AFQT score percentile g, and human capital stock h_1 (uniform marginal). These endowments are drawn from a Gaussian copula.

Students are also endowed with idiosyncratic preferences for individual colleges $\mathcal{U}_q$. These represent flow utilities received while enrolled in any given college q.

2.3 Colleges

Colleges are differentiated by their "quality" $q \in \{1, 2, 3, 4\}$. Each quality group contains one representative college. Colleges of quality 1 correspond to two-year colleges. Students must exit these colleges after two years without earning a degree. All other colleges are four-year colleges where students may earn bachelor's degrees. Students may attend these colleges for up to six years.[2]

[2] We abstract from the option of transferring from two-year to four-year colleges. The main reason is that such transfers are not common in our data.

Colleges differ in their human capital production functions, graduation and dropout rates, and in terms of financial variables, such as college costs. These differences are described in Section 2.5. Higher quality colleges produce more human capital, at least for high ability students, but may also cost more and impose more stringent graduation requirements. This is the main financial trade-off facing students who decide which college to attend.

2.4 Work Phase

It is convenient to describe the life-cycle of a student starting from its last phase, work and retirement. Upon completion of schooling, individuals work from age t_w (the endogenous age after finishing education) to age T_r (physical age 65). Thereafter, workers are retired until they die at age T (physical age 80).

Workers begin their careers endowed with state vector $s_w = (h, k_w, e, t_w)$ (human capital h, assets or debt k_w, education level e, and age t_w). Education e takes on the values HSG for no college, SC for some college without a degree, or CG for college graduates.

Workers solve a simple permanent income problem. Taking the education-specific skill price (w_e) and interest rate (R) as given, they choose the stream of consumption flows $\{c_t\}_{t=t_w}^{T}$ to maximize lifetime utility discounted at rate β. The worker's problem is given by

$$W(s_w) = \max_{\{c_t\}} \sum_{t=t_w}^{T} \beta^{t-t_w} \left[\frac{c_t^{1-\theta}}{1-\theta} + \mathcal{U}_e \right] \tag{1}$$

subject to a lifetime budget constraint that equates the present value of consumption with the present value of labor earnings plus initial assets,

$$\sum_{t=t_w}^{T} R^{t_w - t} c_t = \sum_{t=t_w}^{T_w} R^{t_w - t} w_e h f(t - t_w, e) + k_w. \tag{2}$$

Period utility depends on consumption c_t and the flow utility from leisure and other amenities $\mathcal{U}_e$ associated with jobs typical to education group e. $\theta \geq 0$ is the inverse of the intertemporal elasticity of substitution.

In the lifetime budget constraint, $y \equiv w_e h f(t - t_w, e)$ denotes earnings at age t. $f(\cdot)$ captures how worker productivity varies with experience $(t - t_w)$. We normalize $f(0, e) = 1$.

2.5 College Phase

While enrolled in college, each period unfolds as follows:

1. Students enter the period with state $s = (a, p, g, \mathcal{U}_q, q, h, k, t)$ containing the fixed endowments drawn at high school graduation $(a, p, g, \mathcal{U}_q)$, college quality q, the time varying values of human capital h and assets k, and age t.
2. Students consume and accumulate debt according to the budget constraint

$$c(s) = y(s) + z(s) + Rk[s] - k'(s) - \tau_{total}(s), \tag{3}$$

   From hereon, we write $k[s]$ to denote the k element of s. By contrast, $c(s)$ means that c is a function of s. In the budget constraint $\tau_{total}(s)$ denotes the net net cost of college (tuition minus scholarships or grants), $z(s)$ denotes parental transfers, $k'(s)$ denotes student assets (or debts), and $y(s)$ denotes labor earnings. All financial variables are assumed to depend only on observable student and college characteristics and may therefore be taken directly from the data. Section 2.8 explains this modeling choice.
3. Students enjoy flow utility given by

$$\mathcal{U}_{coll}\left(c, q\right) = \frac{c^{1-\theta}}{1-\theta} + \mathcal{U}_q + \mathcal{U}_{2y} * \mathbb{I}_{q=1}, \tag{4}$$

   where $\mathcal{U}_q$ is a college-specific preference shifter. Students who attend two-year colleges also receive the fixed flow utility $\mathcal{U}_{2y}$ which captures benefits such as living with parents or flexible class schedules.[3]
4. Students accumulate human capital h' (see Section 2.5.1).
5. At the end of the period, students drop out with exogenous probability $\Pr_d(s)$ in which case they start work as college dropouts ($e = SC$) next period. All two year students drop out at the end of year 2. With probability $\Pr_g(s)$ students graduate in which case they start working as college graduates ($e = CG$) next period. Four-year students who have not graduated by the end of year $T_q = 6$ years must drop out. Students who have neither dropped out nor graduated return to college next year.

2.5.1 Learning in College

While enrolled in college, students accumulate human capital according to

$$h' = h + \mathcal{A}(q, a)h^{\zeta}, \tag{5}$$

[3] In our dataset, for 90% of two-year college students, family home is within the 50 miles radius of their college.

where learning productivity is given by

$$\ln \mathcal{A}(q,a) = A_q + \phi_{qq} a + \phi_q \max\left(0,a\right)^2 \mathbb{I}_{q=4} \tag{6}$$

with $A_q, \phi_{qq}, \phi_q \geq 0$. A_q denotes the baseline productivity of college q enjoyed by all students. The remaining terms in equation (6) imply that learning productivity increases with student ability. We impose that ϕ_{qq} is increasing in college quality, so that the productivity gains from upgrading to a better college increase with student ability. In our model, this is the main reason why assigning the "best" students to the "best" colleges maximizes aggregate human capital and earnings.

Finally, we allow for the possibility that high ability students enjoy additional productivity gains from attending the top quality college by setting $\phi_q \geq 0$. We find that this kind of complementarity is needed for the model to match the patterns observed in earnings data.[4]

2.5.2 Value of Studying

The expected value of studying is given by

$$\mathcal{V}\left(s\right) = \mathcal{U}_{coll}\left(c\left(s\right), q[s]\right) + \beta \tilde{\mathcal{V}}\left(s'\right), \tag{7}$$

where $h\left[s'\right] = h'\left(s\right)$ is determined by the human capital technology (5), student debt $k\left[s'\right] = k'\left(s\right)$ is taken from the data, consumption $c\left(s\right)$ is determined by the budget constraint equation (3), and, of course, $t\left[s'\right] = t[s] + 1$. All other elements of s are age invariant. The continuation value is given by

$$\begin{aligned} \tilde{\mathcal{V}}\left(s\right) \quad = \quad & \Pr_d\left(s\right) \mathrm{W}(t[s], h[s], k_w\left(s\right), SC) + \Pr_g\left(s\right) W\left(t[s], h[s], k_w\left(s\right), CG\right) \\ & + \left(1 - \Pr_d\left(s\right) - \Pr_g\left(s\right)\right) \mathcal{V}\left(s\right). \end{aligned} \tag{8}$$

With probability $\Pr_d$, the student drops out and starts work as a college dropout with value $W(., SC)$, defined in equation (1). With probability $\Pr_g$, the student starts work as a college graduate with value $W(., CG)$. With complementary probability, the student remains in college for one more period.

$k_w\left(s\right)$ denotes the worker's assets (or debts) at career start. We assume that a student receives a fixed amount of lifetime transfers, regardless of the college attended or of how long

[4] We have explored models with peer effects where the mean ability of enrolled students affects college productivity. However, for all of the counterfactuals studied in this paper, college's mean student ability levels change very little and therefore peer effects play essentially no role.

the student attends college. While in college, the student receives a portion of this fixed total, $z[s]$. When the student starts their work phase, the remaining transfers are received as a lump-sum, augmenting $k_w(s)$.

The motivation for this assumption is as follows. If students only receive transfers $z[s]$ while in college (and nothing more when they start working), the net cost of college from the student's perspective is tuition minus transfers, $\tau[s] - z[s]$. In the data, this net cost decreases with college quality for high income students. Hence, these students view high quality colleges as cheaper than low quality colleges. This implication strikes us as unreasonable. Our assumption that total transfers are independent of college choice avoids this implication. For the students in our model, the net cost of college is simply tuition $\tau[s]$.

2.6 College Entry Decision

2.6.1 Information Frictions

Our model allows for the possibility that students imperfectly observe college characteristics. We include this information friction for two reasons:

1. Empirical evidence suggests that lack of information may be an important reason why high achieving students, especially those from low income families, choose less selective colleges (Dynarski et al., 2022a).
2. The information friction allows the model to match empirical evidence that college enrollment is highly sensitive to financial incentives. The studies summarized in Page and Scott-Clayton (2016) imply that a \$1,000 increase in annual tuition reduces enrollment by about three to four percentage points. The response is larger for lower income students. In our model, uncertainty about college quality reduces the expected earnings gains from choosing more expensive, higher quality colleges. This uncertainty increases students' willingness to switch colleges in response to changes in tuition costs.

We implement the information friction as follows. Each student is admitted to a subset of colleges $\mathcal{S}$. The admissions decision is based on observable student characteristics as described in section Section 2.7. Students observe the admissions set $\mathcal{S}$ but are uncertain about the human capital productivity, as well as the dropout and graduation probabilities associated with each four year college. All other college characteristics, including financial variables and the student's own preferences $\mathcal{U}_q$, are perfectly observed. Students are also able to identify the two-year college.

For each college in the admitting set $q \in \mathcal{S}$, the student draws a quality signal $\hat{q}(q)$. With probability $\pi(p)$, all signals are accurate so that $\Pr(q|\hat{q}) = 1$ if $q = \hat{q}(q)$ and zero otherwise.

With probability $(1-\pi(p))$, the signals contains no information and the student assigns equal probability to each college in the admitting set so that $\Pr(q|\hat{q}) = 1/n_{\mathcal{S}}$ for each $q \in \mathcal{S}$, where $n_{\mathcal{S}}$ is the number of admitting colleges.

We allow for $\pi(p)$ to depend on parental income because empirical evidence suggests that lack of information affects low income students more than high income students. We assume that students consider only the quality signal when forming beliefs about college quality. In particular, students do not consider financial variables. If they did, the information friction would disappear.

2.6.2 Expected Value of Choosing Signal $\hat{q}$

The expected value of a student who chooses signal $\hat{q}$ is given by

$$\hat{\mathcal{V}}(s,\hat{q}) = \pi(p)\mathcal{V}(\hat{s}(s,\hat{q},\hat{q})) + (1-\pi(p))\sum_{q^*\in\mathcal{S}}\mathcal{V}(\hat{s}(s,q^*,\hat{q}))/n_{\mathcal{S}}, \tag{9}$$

where $\hat{s}(s,q^*,\hat{q})$ denote the perceived state of a student with state s (ignoring the implied college quality) who chooses the college with signal $\hat{q}$ but ends up with the productivity of college q^*.

With probability $\pi(p)$ the student observes the true quality and starts college $\hat{q}$ with state $\hat{s}(s,\hat{q},\hat{q}) = (a,p,g,\mathcal{U}_{\hat{q}},\hat{q},h,k,t)$. With complementary probability, the student expects to start college with finances determined by $\hat{q}$ but with college quality determined by a randomly drawn quality q^*.

2.6.3 College Entry Decision

After high school graduation, students either choose one of the colleges they are admitted to or they begin work with education level HSG. Students choose the option that yields the highest expected value:

$$\hat{q}(q) = \arg\max\{W(s_w), \{\hat{\mathcal{V}}(s,\hat{q}(q))\}_{\hat{q}\in\mathcal{S}}\} \tag{10}$$

where the value of working as a high school graduate is obtained from (1). The true college quality implied by the chosen signal $\hat{q}$ is revealed when the student enters college.

2.7 College Admissions

Our model of admissions is broadly based on Hendricks et al. (2021). It captures a number of desirable features in a tractable way:

1. Selective colleges are capacity constrained and reject academically qualified applicants.[5]
2. In addition to test scores and grades, college admissions consider other indicators of college preparation, such as extracurricular activities or AP exam scores. For given measured ability (e.g., test scores), higher income students likely perform better according to these indicators.[6]
3. For given test scores, admission rates to selective colleges rise with family income. The differences in admission rates are partly due to the fact that low income students do not apply to selective colleges (perhaps because they expect not to be admitted; see Marto and Wittman (2024)). As we discuss in section Section 2.8, we think of the admissions process in the model as encompassing frictions in the application process that prevent low income students from applying to selective colleges.

We model admissions as follows. Each student's endowments imply an admissions score z. Colleges aim to attract students with high scores, but are subject to capacity constraints. Each college therefore admits all students with scores above a cutoff, $z \geq \bar{z}_q$. The cutoffs are set such that all four-year colleges are full. Two year colleges admit all students and face no capacity constraints.

Students choose colleges sequentially in order of their admission scores. The student with the highest z chooses first and is admitted to all colleges. The student with the second highest z chooses next, and so on. As students enroll, college seats are filled. Once a college reaches its enrollment capacity, it no longer admits students. The last student admitted determines the cutoff $\bar{z}_q$. Students with $z < \bar{z}_q$ do not have college q in their admissions set $\mathcal{S}$.

The admissions score z is a linear combination of test score g percentiles and human capital h_1 percentiles (with weights $\beta_g, \beta_h \geq 0$). The functional form captures the idea that admissions officers consider not only academic achievement (test scores or high school grades), but also other indicators of college preparation, such as extracurricular activities or AP courses taken. The human capital endowment h_1 proxies for these indicators, which are correlated with student ability and parental background.

Students with low admissions scores are rationed out of selective colleges. This is one reason

[5] Carnevale and Rose (2004, p. 6) conclude that selective colleges "could in fact admit far greater numbers of low-income students, including low-income minority students, who could handle the work."

[6] Empirical evidence about the relationship between applicant credentials and parental income is scarce. Alvero et al. (2021) show that the quality of college essays varies with parental income.

for undermatch, especially for low income students who typically have low human capital endowments at high school graduation.

The sequential college choice algorithm of our model avoids the substantial complications and loss of tractability that would arise in models with student applications (Chade et al., 2014; Fu, 2014) or two-sided matching (Epple et al. 2006).

2.8 Discussion of Modeling Choices

2.8.1 Exogenous Dropout Rates

Since the literature has not come to a consensus about the main reasons why students drop out, it would be challenging to model dropout decisions in a compelling way.[7] We therefore treat dropping out as a response to unobserved shocks that we do not model.

One advantage is that the model is able to replicate how empirical dropout rates vary with college quality and observable student characteristics. Accurately capturing the financial returns to college (quality) is important for understanding the implications of reallocating students across colleges.

One drawback is that college appears riskier when dropping out is a shock compared with the case where dropping out is a choice. The option of dropping out limits the downside risk of trying college when outcomes are uncertain.[8]

2.8.2 Exogenous Consumption and Borrowing

We also do not model consumption-savings decisions while in college. Instead, we assume that all financial variables (college costs, transfers, and borrowing) only depend on observables and may therefore be directly taken from the data.

In part, we make this choice to ensure that the model correctly captures the observed financials of students with different backgrounds who are enrolled in colleges of different qualities. In part, the choice is due to data limitations. We lack evidence about how much students would have to pay for colleges that they do not attend in the data. Similarly, we lack evidence on the extent to which parental transfers would cover the additional costs incurred by attending a better college.

---

[7] Bound and Turner (2011, p. 605) conclude: "In hypothesizing about why students leave college without receiving a degree, the research literature has posited many ideas ranging from learning about own ability to clear 'mistakes' in the utilization of financial aid or the navigation of complicated collegiate requirements."

[8] Since we do not consider counterfactual experiments that change students' incentives to drop out, our results are not affected by the Lucas critique.

One drawback is that our model may understate the importance of borrowing constraints. If some students in the data fail to attend selective colleges because of unobserved financial tightness (e.g., parents are not willing to make substantial transfers to pay for college), our model misses that constraint. Whether financial constraints prevent substantial numbers of students from entering college or choosing selective colleges remains controversial in the literature.[9]

It is worth noting that, in our data, most student borrowing is far from federal student debt limits. About half of all four-year college entrants do not borrow at all. These numbers suggest that, consistent with our model's implications, financial constraints may not be of first order importance for college choice.

2.8.3 College Admissions

Our model implies that many students are rationed out of high quality colleges. Especially low income students are often not admitted, even if they have high ability. Not being admitted to selective colleges is a key reason why model students are undermatched. How plausible is this model implication?

Empirical evidence suggests that admissions at selective colleges indeed favor high income students. One reason is that high income students apply with strong non-academic credentials, such as extracurricular activities or application essays (Carnevale and Rose, 2004; Alvero et al., 2021).

"Top percent" programs offer direct evidence that admissions prevent low income, high ability students from attending selective colleges. California's "Eligibility in the Local Context" program grants admissions to some selective public universities for high school students with GPAs in the top four percent of their school peers. Bleemer (2024) finds that the program increased enrollment of eligible students at those universities by nearly one third (ten percentage points). Similarly, Black et al. (2023a) and Kapor (2024) find that Texas's "Top Ten Percent Rule" nearly doubled enrollment by treated students at UT Austin (the flagship university). These large enrollment responses suggest that a substantial number of high ability, mostly low income students fail to attend selective colleges because they are not admitted.

Many colleges adjust their admissions standards for the differences in opportunities enjoyed by students of different income levels. However, it is not clear that these adjustments eliminate

[9] One reason why "the literature has yet to reach a consensus on the extent to which constraints discourage youth for recent cohorts" Lochner and Monge-Naranjo (2011, p. 237) is that exogenous variation in credit availability is hard to find, given that most U.S. college students have had access to federal student loans for a long time (Dynarski et al., 2022b).

the admissions advantages of rich students (Bowen et al., 2005). It is therefore likely that low income students are admitted at lower rates than high income students with similar ability levels. But, since ability is not directly observable, clear evidence is lacking.

Low income students face additional frictions that prevent many from applying to selective colleges. Some are poorly informed about college costs (Dynarski et al., 2022a). Others may expect that they will either not be admitted at all or only with an unattractive financial package (Kapor, 2024; Marto and Wittman, 2024). Some high schools provide limited support for navigating a complex application process. (Hoxby and Avery, 2013; Roderick et al., 2011).[10] Even after being admitted, a significant number of students fail to enroll (Castleman and Page, 2015).

We model some of the frictions facing lower income students, including financial constraints, the lack of good information about colleges, and idiosyncratic preferences for specific colleges. The remaining unmodeled frictions are absorbed into the model's admissions process. We therefore think about model admissions broadly as encompassing not only college admissions but also unmodeled frictions that prevent students from applying to selective colleges.

3 Calibration

This section outlines the calibration strategy and summarizes the model fit. Details are relegated to Appendix A.

3.1 Data

Our main data source is the 1997 cohort of the National Longitudinal Survey of Youth (Bureau of Labor Statistics; US Department of Labor, 2002). The NLSY97 is an ongoing panel dataset that surveys youth born between 1980 and 1984. Leukhina (2023) describes the data in detail.

3.1.1 College Quality Groups

We distinguish between four college quality levels. Quality is measured by mean freshman SAT scores.[11] Quality group 1 comprises community colleges offering an associate degree in

[10] Dynarski et al. (2022a) survey the evidence supporting non-financial barriers to college enrollment.

[11] How to measure college "quality" is debated in the literature. The survey by Lovenheim and Smith (2023, Section 4.2) concludes that "[m]ore often than not, approaches using these various measures find consistent results" (p. 39). Other studies that classify colleges based on mean SAT scores include Bowen et al. (2009) and Dillon and Smith (2017). Given that our quality categories are broad, it is unlikely that other

Table 1: College Quality Summary Statistics

		Quality			
	All	**1**	**2**	**3**	**4**
Mean AFQT percentile	63	50	61	73	83
Graduation rate (cond.)	0.44	-	0.53	0.74	0.85
Mean tuition	6,704	2,060	6,001	7,349	12,991
Mean net college cost	2,118	795	1,153	2,692	5,590
SAT cutoff	-	-	-	1,033	1,136
Sample size	1,495	586	342	308	259

Note: The table summarizes freshmen characteristics by college quality. Quality 1 comprises 2-year colleges. Quality categories 2 to 4 refer to 4-year institutions, ranked from least to most selective. "Graduation rate (cond.)" is the fraction of freshmen who graduate within six years. "SAT cutoff" is the lowest freshmen SAT score in the quality group. College SAT scores are the average of the 25th and the 75th percentiles.

general education. Quality groups 2 through 4 represent four year colleges and universities that grant bachelor's degrees. Each group of four year colleges has approximately equal freshman enrollment. Quality group 4 comprises Ivy-league and selective private schools, most flagship universities and many other selective public universities, such as UNC Chapel Hill or Iowa State University.[12] Quality 3 includes most of the remaining flagship universities and state schools, such as the University of Arizona or Michigan State University. Quality 2 colleges include the least selective public schools and many for-profit private colleges, such as San Diego State University, East Carolina University, or Stillman College. Table 1 shows summary statistics.

3.2 Fixed Parameters and Assumptions

This section summarizes the model parameters that are fixed based on outside evidence. The model period is one year.

Preferences: We set the curvature of utility from consumption to $\theta = 1.5$ and fix the discount factor at $\beta = 0.96$.

commonly used quality definition would substantially change our findings.

[12] It would be desirable to create a separate quality group for the highly selective colleges that receive much attention in the public discussion. Unfortunately, our sample sizes are too small to obtain reliable results for this analysis.

Worker experience profiles $f(x, e)$ are estimated using the NLSY's longitudinal earnings histories. Since we only observe roughly the first fifteen years of workers' careers, we extend the profiles by splicing on education-specific experience profiles estimated in Rupert and Zanella (2015). The resulting profiles are shown in Figure 8.

Education-specific skill prices are calibrated. We assume that college graduates enjoy a sheepskin effect: $w_{CG} \geq w_{SC}$. The skill price for dropouts is the same as for high school graduates: $w_{SC} = w_{HSG}$. This assumption avoids artificial wage increases for students who attend college for only short periods without learning much.

The gross interest rate is fixed at $R = 1.04$.

3.2.1 Colleges

Capacities: We set college capacities for four year colleges to their empirical freshmen enrollment levels. The two year college has unlimited capacity.

Dropout and graduation rates: The probabilities of dropping out of college, $\Pr_d(s)$, and of graduating from college, $\Pr_g(s)$, are both functions of student ability percentiles $\hat{a}$ and parental income percentiles p. The functions differ across colleges but not over time. Students can only graduate after attending a four year college for at least 3 years. This simple specification results in a good empirical fit.

For graduation, we assume that $\Pr_g(s) = (\gamma_{1,q} + \gamma_{2,q}\hat{a}) \times (1 - \gamma_{3,q} \times (1-p))$ where $\gamma_{1,q}, \gamma_{2,q} \geq 0$ and $0 \leq \gamma_{3,q} \leq 1$. For given parental income, the probability of graduating is linear in ability percentile. Students with higher parental incomes are more likely to graduate. Similarly, for dropping out, we assume that $\Pr_d(s) = (\gamma_{4,q} - \gamma_{5,q}\hat{a}) \times (1 - \gamma_{6,q} \times p)$ where $\gamma_{4,q}, \gamma_{5,q} \geq 0$ and $0 \leq \gamma_{6,q} \leq 1$. Students with higher ability or higher parental income are less likely to drop out. All probabilities are truncated into the unit interval.

College finances: We directly estimate all financial variables from the data. We assume that most financials do not differ across years for two reasons. First, sample sizes get smaller over time as students drop out, making it difficult to estimate time variation. Second, financial variables in later years may be affected by selection if, for example, students with limited resources drop out at high rates. The details are as follows:

- College costs: The annual net cost of attending college $\tau_{total}(s)$ is the sum of an observed cost $\tau(s)$ and an additional (calibrated) unobserved cost τ_{4y} that is paid by all four-year students. The observable cost is estimated by regressing tuition charges, net of grants and scholarships, on family income, test scores, and college quality (see Table 4). As expected, observed college costs increase with college quality and parental income, but decline with test scores. The cost of attending a four-year college helps the model match the fact that higher income students are more likely to attend such colleges.

- Parental transfers: We set parental transfers to their observed means for each combination of family income quartile and college quality (see Table 5).
- Student debt: We find that, for given college quality, debt varies little with AFQT scores or parental incomes. We therefore set debt for all students to the estimated means by college quality and year (see Table 6). We assume that annual borrowing stays constant after year four, which is the last year for which we have enough observations to estimate debt with reasonable precision. In our data, students rarely borrow large amounts. At the end of their fourth year in college, mean debt is just over \$10,000 and almost half of students have no debt at all.
- Student earnings: In our data, student earnings vary little with parental incomes or student test scores. We therefore set student earnings to their estimated means for all students in a given college. Earnings are similar for all four-year colleges, but higher for two-year colleges.[13]

3.3 Calibration Strategy

We calibrate 43 model parameters by minimizing a weighted sum of squared deviations between data moments and simulated model moments. This section provides a summary with the details relegated to Appendix A.

The calibrated parameters include:

- Endowment correlations and marginal distributions.
- Preferences: $\mathcal{U}_{2y}$, $\mathcal{U}_e$, and the range of $\mathcal{U}_q$, which is drawn from a uniform distribution with mean zero.
- Human capital production functions: A_q, ϕ_{qq}, ϕ_q, ζ.
- Graduation probabilities $(\gamma_{1,q}, \gamma_{2,q}, \gamma_{3,q})$ and dropout probabilities $(\gamma_{4,q}, \gamma_{5,q}, \gamma_{6,q})$.
- Information frictions: $\pi(p)$ for each parental income quartile.
- Admissions: the parameters that determine admissions scores (β_h and β_g), and admissions cutoffs $\bar{z}_q$.
- Skill prices: w_e.

Many of the calibrated parameters have no clear observable proxies. The calibration therefore requires a large number of data moments to pin down all parameters. The target moments

[13] The dollar values by college quality are: $q1$: \$8,100, $q2$: \$5,442, $q3$: \$4,651, $q4$: \$4,430.

may be summarized as follows:

1. High school graduate endowments:

   The fraction of high school graduates in each parental income and AFQT quartile.

2. College enrollment patterns:

   College entry rates by quality, parental income and AFQT quartile.

   Mean freshman AFQT percentiles by college quality.

   Freshmen enrollments by quality.

3. College graduation rates:

   Fraction of entrants who graduate by quality, parental income and AFQT quartile.

   Average time to graduation by quality or AFQT quartile.

4. College dropout rates:

   Average time to dropout by quality or AFQT quartile.

   Cumulative dropout rates after year two by quality and AFQT quartile.

5. Worker earnings:

   Regressions of log earnings (net of experience effects) on education, college quality, and AFQT quartile.

In addition, we target two quasi-experimental data moments:

1. The response of college enrollment to changes in tuition:

   Based on the literature survey by Dynarski et al. (2022b), we target an enrollment change of 3.5 percentage points per $1,000 annual change in tuition. This data moment is important for identifying the scale of idiosyncratic college preferences $\mathcal{U}_q$. When college preferences are highly dispersed, college enrollment is insensitive to financial incentives.

2. The effect of providing information about college quality to high AFQT, low income students:

   Our intervention approximates that of Hoxby et al. (2013) who sent information about potential colleges to high school graduates with parental incomes in the lowest tercile and test scores in the top decile. Their intervention treats only about 1.5 percent of high school graduates. This fraction is too small to obtain precise results from our simulated $10,000$ student types. We therefore treat students in the lowest half of the parental income distribution with test scores in the top quintile. Treated students are given full information ($\pi = 1$) about college quality. Based on Hoxby et al. (2013), we target an increase in the college entry rate of 5.3 percentage points. This data moment mainly helps to identify $\pi(p)$.

In some cases we "slice" the same data in different ways that may appear redundant, but are in fact important to pin down certain parameter values. For example, we run two sets of earnings regressions. One includes all workers. The second focuses on college graduates. The main purpose of the second regression is to estimate complementarities (interactions) between high AFQT students and top colleges.

The calibrated parameter values are shown in Appendix A. The model implies that AFQT scores and ability levels are highly correlated.[14] The high correlation simplifies the interpretation of the findings.

3.4 Model Fit

Overall, the calibrated model fits most of the targeted moments well. In this section, we highlight a few of the moments that are relevant for the discussion of the results in Section 4. Appendix D shows the model fit for the remaining target moments.

Table 2 shows a regression of college graduate log earnings (net of experience effects) on AFQT and quality dummies and their interactions. The key implication is that the wage "gains" from attending top quality colleges mostly accrue to top AFQT students. Through the lens of the model, this finding suggests a form of *complementarity* between student ability and college quality. Specifically, the model implies that learning productivity is especially high for high ability students who attend the top college (see Figure 1). This property plays an important role for understanding the implications of public finance policies. For both intergenerational mobility and aggregate earnings, giving high ability students access to top colleges is key. This complementarity coupled with the fact that a lot of high ability students are rationed out of top quality colleges is what can potentially drive aggregate earnings gains from reallocating funds towards high quality colleges.

Figure 2 shows the joint distribution of AFQT scores and parental incomes. Note that a substantial fraction of low income students have high AFQT scores.

Figure 3 shows the fraction of students in each AFQT quartile that choose each college. Most high AFQT students do not attend top quality colleges. This well-known observation is labeled "undermatch" in the literature. Undermatch is most prevalent among low income students.[15] As shown in Figure 4, low income students rarely enroll in top quality colleges.

[14] The correlation between AFQT and ability is mainly identified by the AFQT coefficients in the earnings regressions. If the model is recalibrated while fixing this correlation at lower values, these coefficients are smaller than in the data. In addition, the model fails to replicate some of the AFQT gradients in college entry rates.

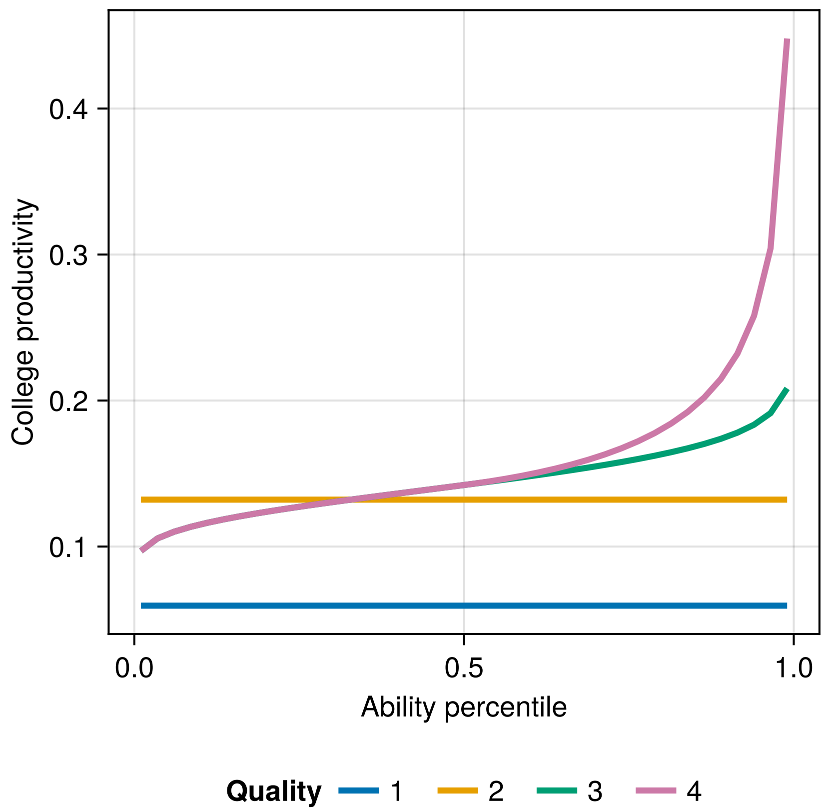
[15] The literature employs various definitions of undermatch, but all aim to measure the fraction of qualified

Table 2: Earnings Regressions for College Graduates

Regressor	**Data**	**Model**
AFQT 2	0.0217	0.0109
	(0.0563)	
AFQT 3	0.0365	0.0334
	(0.0561)	
AFQT 4	0.004266	0.0345
	(0.0710)	
AFQT4-Qual3	0.0641	0.0641
	(0.0709)	
AFQT4-Qual4	0.207	0.203
	(0.0858)	
Quality 3	0.0534	0.0534
	(0.0412)	
Quality 4	0.0793	0.0857
	(0.0548)	
Constant	2.94	2.91
	(0.0528)	

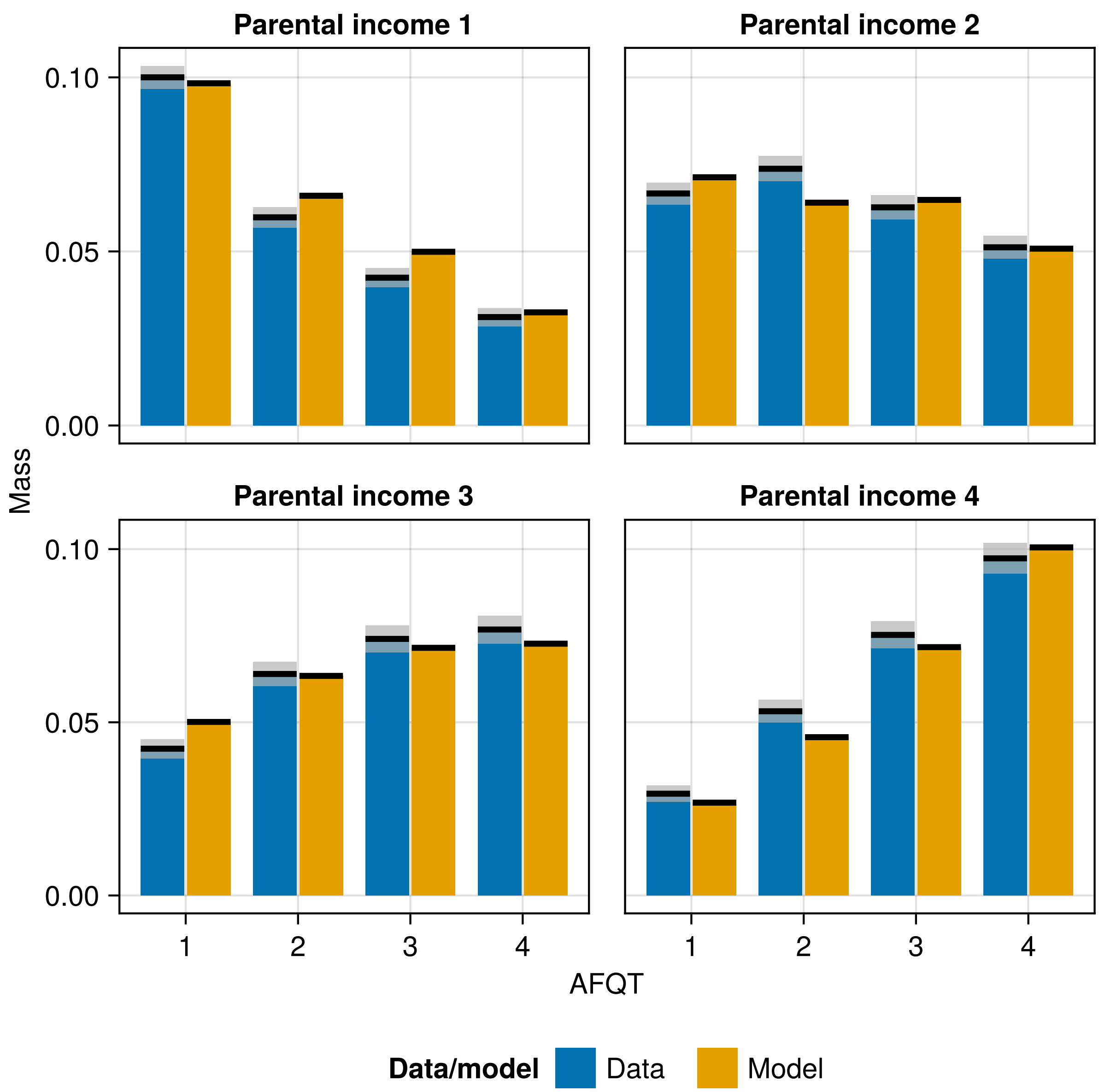
Note: The table shows the coefficients and standard errors (in parentheses) of an earnings regression for college graduates. The dependent variable is log earnings net of experience effects. The regressors include dummies for AFQT quartiles, college quality groups, and selected interactions.

Figure 1: Learning Productivities



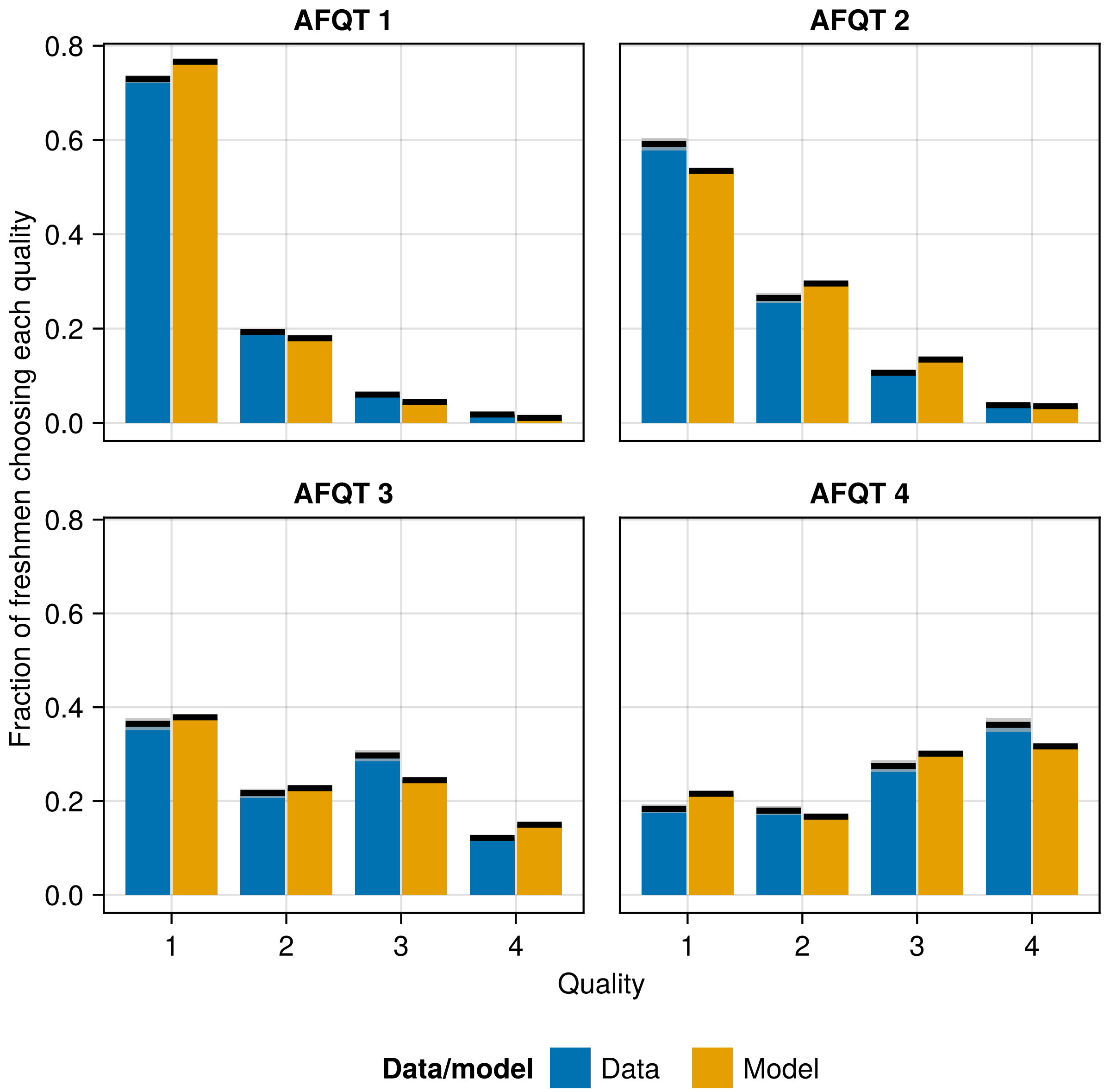
Note: The Figure shows learning productivity $\mathcal{A}(q, a)$ as a function of ability percentile for each college.

Figure 2: Joint Distribution of AFQT and Parental Income



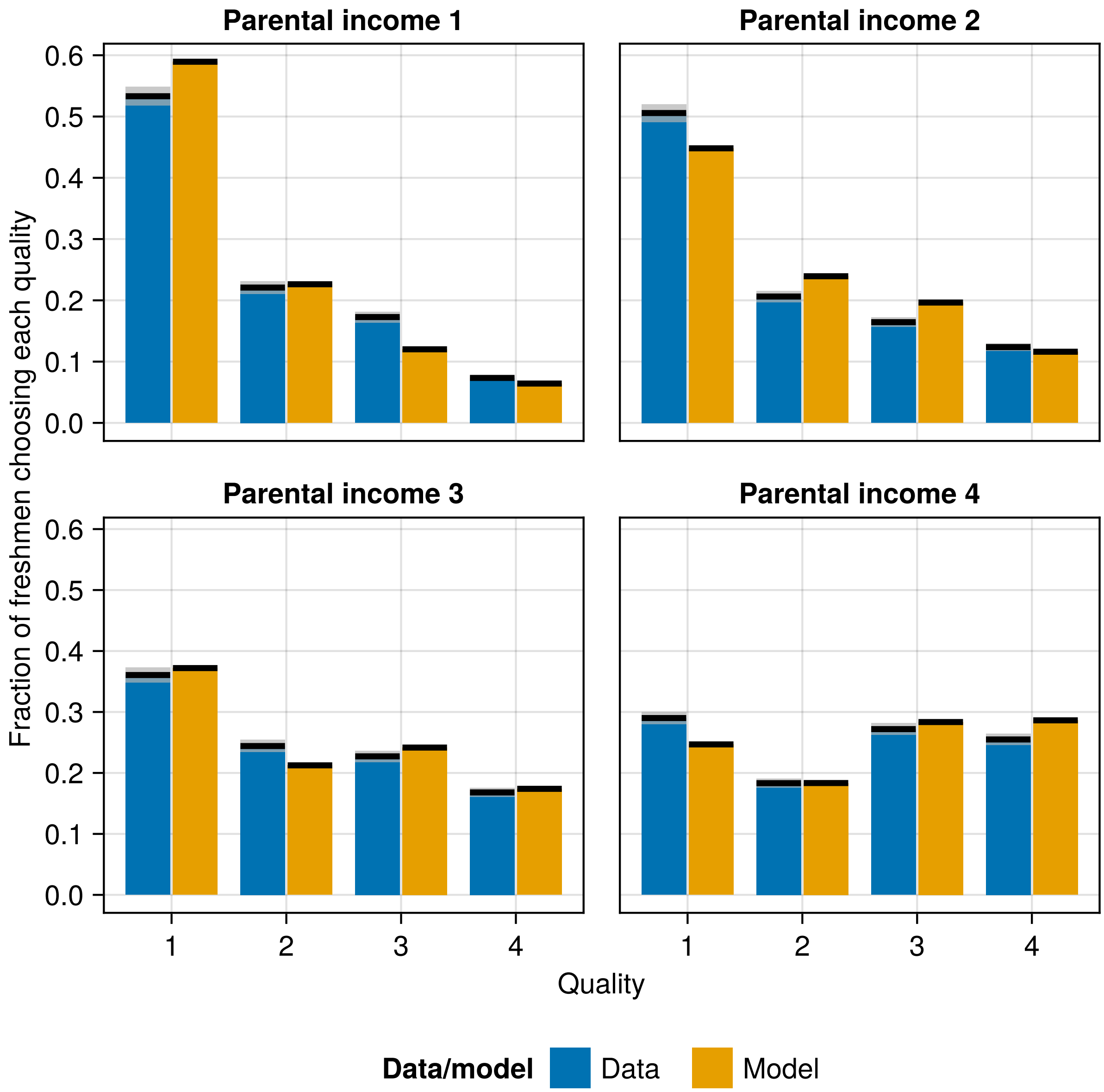
Note: The Figure shows the mass of high school graduates in each AFQT and parental income quartile.

Figure 3: Distribution of Freshmen



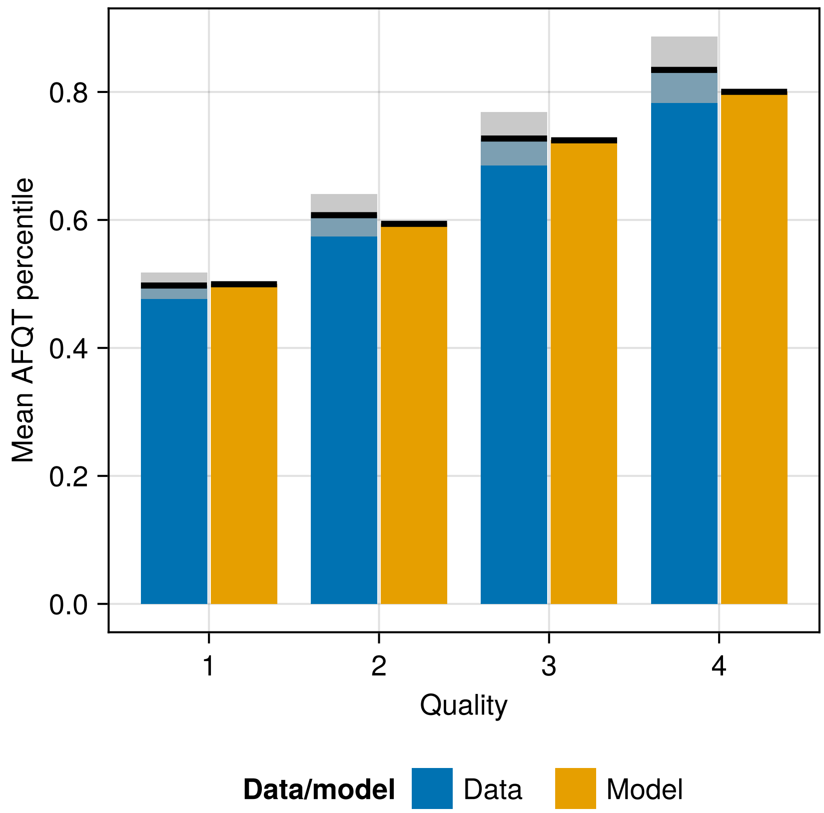
Note: The Figure shows the fraction of college freshmen in each AFQT quartile that choose each college. The fractions sum to one for each AFQT quartile.

Figure 4: Distribution of Freshmen



Note: The Figure shows the fraction of freshmen in each parental income quartile who choose each college. The fractions sum to one for each parental income quartile.

Figure 5: Mean AFQT Percentiles by College Quality



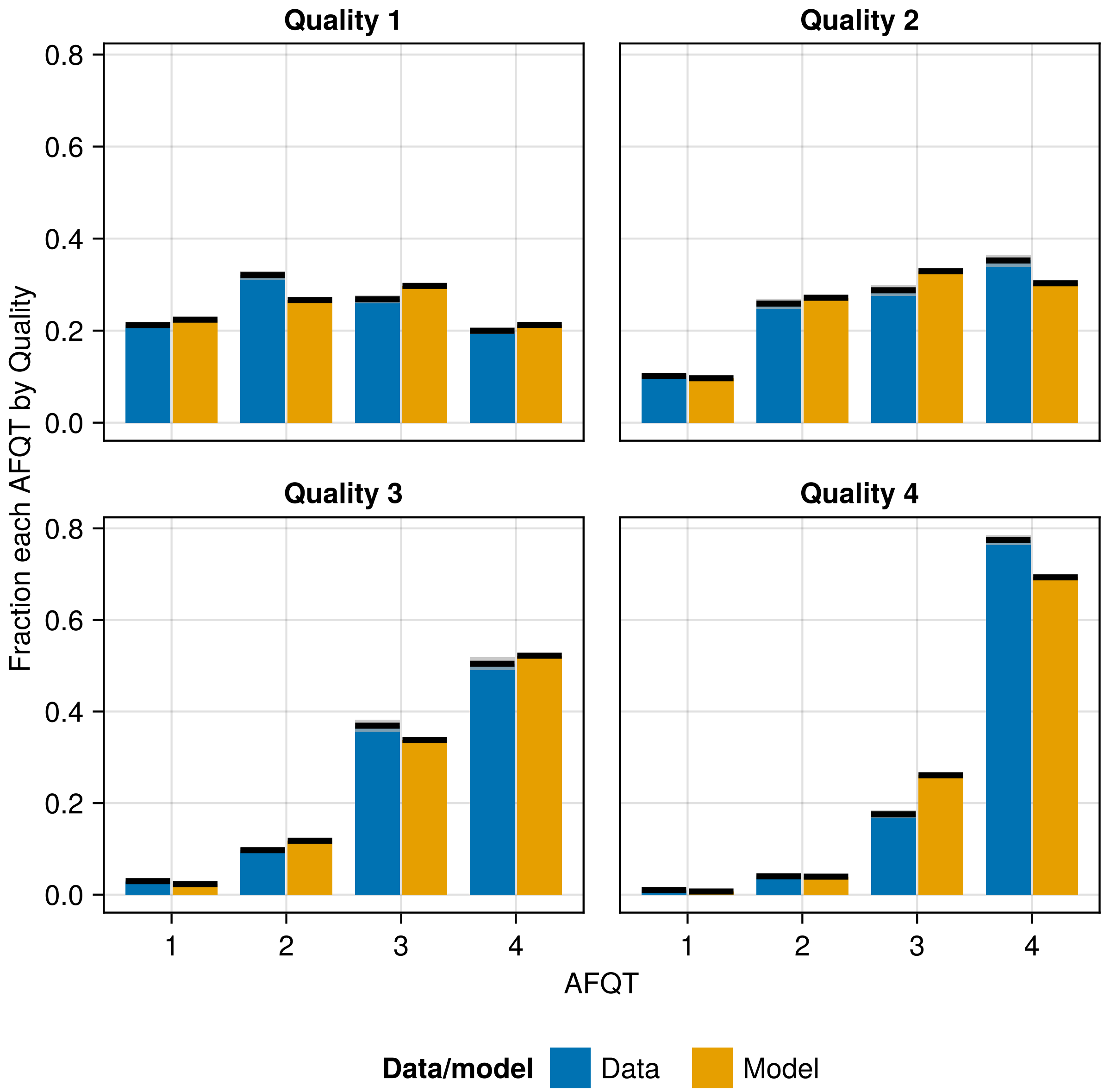
It follows that there is sizable pool of students with high measured ability that appropriate policies could potentially attract to selective colleges.

Even though the matching between student ability and college quality is highly imperfect, mean student AFQT scores are much higher for more selective colleges (see Figure 5). While low quality colleges are attended by students of all AFQT levels ("undermatch"), high quality colleges effectively ration out low AFQT students (see Figure 6).

Taken together, these observations suggest that expanding access to selective colleges may lead to large financial and welfare gains. There is a substantial pool of (low income) high ability students (as proxied for by AFQT scores; Figure 2), most of whom do not attend the top college (Figure 3). If expanding capacity of high quality colleges were to attract these students, these students would enjoy substantial earnings gains when they upgrade to the top

students who fail to enroll in appropriately selective colleges. For evidence on undermatch, see Bowen et al. (2009) or Dillon and Smith (2017).

Figure 6: AFQT Distribution by College Quality



Note: For each college, the Figure shows the fraction of freshmen in each AFQT quartile.

college (Table 2), leading to large aggregate gains in human capital. The question is whether or not the extra financial gains justify the increased expense of creating these seats.

4 Results

4.1 The Public Finance Policy Experiment

We study the implications of a budget-neutral reallocation of funds from college quality 2 to quality 4. The reallocated funds increase the capacity of quality 4 by 10 percent. We use IPEDS data on college-level expenditures and student enrollment to calculate average instruction-related (broadly defined) expenditures per student in each type of college. Appendix Section D.4 provides the description of each category of college-level expenditures that we include in our aggregate measure of instruction-related expenditures. Conceptually, we want to include all categories that would allow for a small expansion of the student body without requiring the change in infrastructure.

We find that creating an additional spot in a quality 4 college costs about $12,500, while creating a spot in a quality 2 college costs about $5,000. Thus, for every seat created in high quality colleges, a little over two seats are lost in low quality colleges. Thus, the overall college capacity declines.

4.2 Results

This experiment leaves (endogenous) enrollment in quality 1 and 3 colleges nearly unchanged, which simplifies the interpretation of results. In other words, while enrollment in quality 4 colleges increases by 10 percent, the lost seats in quality 2 colleges (over 20 percent of its benchmark enrollment) roughly equal the number of students who opt out of college entry due to reduced capacity of four-year colleges.

Figure Figure 7 shows the mean ability levels by college quality in the benchmark model and in the policy experiment. The mean ability level falls in quality 4 colleges as more students gain access. The marginal entrant are of lower ability, on average, than students attending quality 4 colleges in the benchmark. This is because the admissions index positively correlates with student ability. On the contrary, mean student ability rises in quality 2 colleges as students losing access to four-year schools tend to be of lower ability.

Table 3 summarizes aggregate implications of the public finance policy experiment.

The overall decline in college entry (from .571 to .565) is driven by the reduction of entry by students in the 2nd ability quartile who either do not wish to attend quality 4 colleges or get

Figure 7: Student Ability by Quality

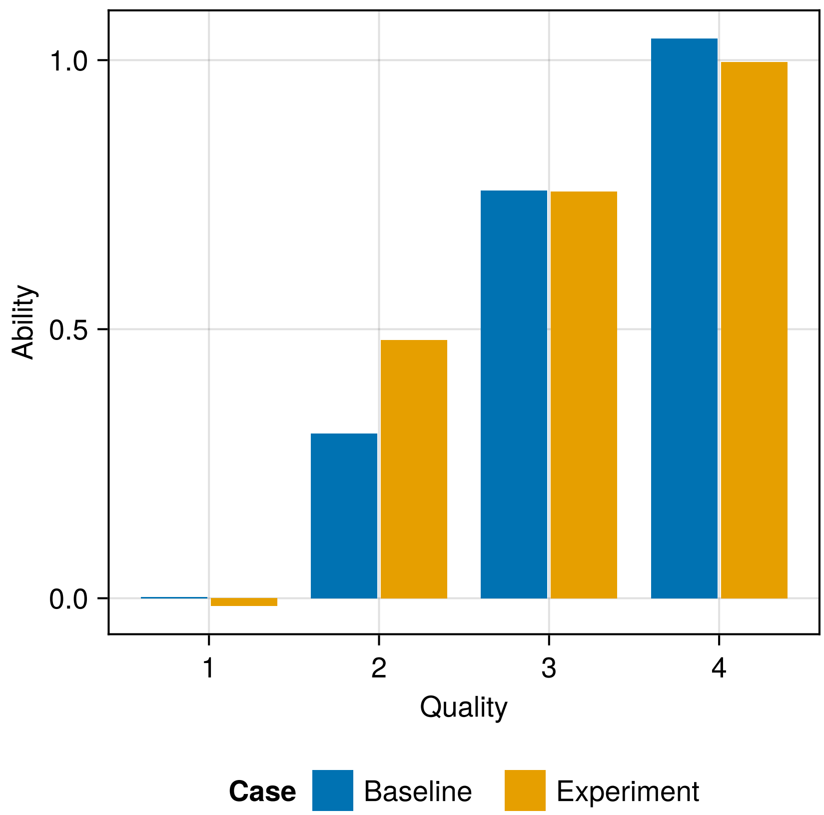


Table 3: Aggregate Outcomes

	Base	Policy
Aggregate outcomes		
Mean log lifetime earnings	6.302	-0.004
Lifetime earnings 90 / 10 gap	94.0	+0.5
Entry rate	57.1	-1.5
Graduation rate (cond.)	41.5	+0.4
Frac high ability in low quality college	38.0	37.2
Intergenerational mobility		
Corr lifetime earnings pct / parental inc pct	56.2	+1.8
Probability quartile 1 to 4	7.1	-1.7
Probability quartile 4, rich vs poor	40.8	+3.1
Lifetime earnings gap by Parental Inc	32.2	+1.6

rationed out.

Graduation rates rise very little (from .415 to .419) as students reallocate from quality 2 to quality 4 colleges. This is because even though a student of fixed ability (for high enough ability levels) faces higher graduation rates in quality 4 schools, the average ability declined. Interestingly, the number of college graduates is roughly unchanged. The product of entry rate and graduation rate in the benchmark model ($.571 * .415 \approx .237$) is approximately the same under the policy experiment ($.565 * .419$). This means that the number of lost seats is approximately accounted for by the reduction in the number of dropouts.

As a result, the loss of aggregate earnings due to exit stems from the loss of dropout premia. While the earnings of college graduates increase – as students reallocate to quality 4 colleges which feature a superior human capital accumulation technology – the exit-driven loss of dropout premia dominates the impact on aggregate earnings. We find that the aggregate lifetime earnings decrease by 0.4 percentage points.

These results are quite striking! Even though we find that human capital accumulation technology and graduation rates are substantially superior in quality 4 colleges, especially for high ability students, and even though we uncover a large supply of high ability students that are rationed out of quality 4 colleges and stand to gain a lot from increased access, we find that reallocting funding from low quality colleges to high quality colleges is not worth it. It does not increase the total number of college graduates and fails to raise earnings. Creating seats in better quality schools does not justify the losses associated with shutting down many more low quality seats.

Likewise, we find that our policy experiment leads to a decline in intergenerational income mobility. While those who upgrade college substantially benefit from the reform, those who are forced to exit college earn less. As a first order approximation, the losers are those who would have benefited from going to low quality college for a couple of years before dropping out.

5 Conclusion

We find that reallocating a fixed amount of public funds from low quality to high quality four-year colleges leads to a decline of both aggregate earnings and upward income mobility. The intuition is as follows. For every seat created in high quality colleges, a little over two seats are lost in low quality colleges – because per-student instruction-related expenditures are substantially lower at these schools. Fewer students enter college because of the overall loss of capacity. The number of college graduates, however, remains roughly unchanged because graduation rates increase with greater access to high quality schools. While the graduates do earn more, consistent with our initial intuition, the loss of access to lower quality seats implies fewer dropouts and therefore lower aggregate earnings.

Thus, the observed expansion of funding directed at lower quality institutions makes sense from the efficiency standpoint.

References

ALVERO, A., S. GIEBEL, B. GEBRE-MEDHIN, A. L. ANTONIO, M. L. STEVENS, AND B. W. DOMINGUE (2021): "Essay content and style are strongly related to household income and SAT scores: Evidence from 60,000 undergraduate applications," *Science advances*, 7, eabi9031.

ARMSTRONG, L. T., ELIZABETH A.; HAMILTON (2013): *Paying for the Party: How College Maintains Inequality*: Harvard University Press.

BASTEDO, M. N., AND O. JAQUETTE (2011): "Running in place: Low-income students and the dynamics of higher education stratification," *Educational Evaluation and Policy Analysis*, 33, 318–339.

BLACK, S. E., J. T. DENNING, AND J. ROTHSTEIN (2023a): "Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes," *American Economic Journal: Applied Economics*, 15, 26–67.

——— (2023b): "Winners and Losers? The Effect of Gaining and Losing Access to Selective Colleges on Education and Labor Market Outcomes," *American Economic Journal: Applied Economics*, 15, 26–67, 10.1257/app.20200137.

BLEEMER, Z. (2024): "Top Percent Policies and the Return to Postsecondary Selectivity."

BOUND, J., AND S. TURNER (2011): "Dropouts and diplomas: The divergence in collegiate outcomes," *Handbook of the Economics of Education*, 4, 573–613.

BOWEN, W. G., M. M. CHINGOS, AND M. S. MCPHERSON (2009): *Crossing the Finish Line: Completing College at America's Public Universities*: Princeton University Press.

BOWEN, W. G., M. A. KURZWEIL, E. M. TOBIN, AND S. C. PICHLER (2005): *Equity and excellence in American higher education*: University of Virginia Press.

BUREAU OF LABOR STATISTICS; US DEPARTMENT OF LABOR (2002): "National Longitudinal Survey of Youth 1979 cohort," Produced and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH.

CARNEVALE, A. P., AND S. ROSE (2004): "Socioeconomic status, race/ethnicity, and selective college admissions."

CASTLEMAN, B. L., AND L. C. PAGE (2015): "Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates?" *Journal of Economic Behavior & Organization*, 115, 144–160.

Chade, H., G. Lewis, and L. Smith (2014): "Student portfolios and the college admissions problem," *Review of Economic Studies*, 81, 971–1002.

Dillon, E. W., and J. A. Smith (2017): "Determinants of the Match between Student Ability and College Quality," *Journal of Labor Economics*, 35, 45–66, 10.1086/687523.

Dynarski, S., A. Nurshatayeva, L. C. Page, and J. Scott-Clayton (2022a): "Addressing Non-Financial Barriers to College Access and Success: Evidence and Policy Implications," Working Paper 30054, National Bureau of Economic Research.

Dynarski, S., L. C. Page, and J. Scott-Clayton (2022b): "College costs, financial aid, and student decisions,"Technical report, National Bureau of Economic Research.

Epple, D., R. Romano, and H. Sieg (2006): "Admission, tuition, and financial aid policies in the market for higher education," *Econometrica*, 74, 885–928.

Fu, C. (2014): "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market," *Journal of Political Economy*, 122, 225–281, 10.1086/675503.

Hendricks, L., C. Herrington, and T. Schoellman (2021): "College Quality and Attendance Patterns: A Long-Run View," *American Economic Journal: Macroeconomics*, 13, 184–215, https://www.aeaweb.org/articles?id=10.1257/mac.20190154.

Hoekstra, M. (2009): "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *Review of Economics and Statistics*, 91, 717–724, 10.1162/rest.91.4.717.

——— (2020): "Returns to education quality," in *The Economics of Education*: Elsevier, 65–73.

Hoxby, C., and C. Avery (2013): "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students," *Brookings Papers on Economic Activity*, 1, https://www.questia.com/library/journal/1G1-351948040/the-missing-one-offs-the-hidden-supply-of-high-achieving, Publisher: Brookings Institution.

Hoxby, C. M. (2009): "The changing selectivity of American colleges," *The Journal of Economic Perspectives*, 23, 95–118.

Hoxby, C., S. Turner et al. (2013): "Expanding college opportunities for high-achieving, low income students," *Stanford Institute for Economic Policy Research Discussion Paper*, 12, 7.

Kapor, A. (2024): "Transparency and Percent Plans,"Technical report, National Bureau of Economic Research.

KINSLER, J., AND R. PAVAN (2011): "Family income and higher education choices: The importance of accounting for college quality," *Journal of human capital*, 5, 453–477.

LEUKHINA, O. (2023): "The Changing Role of Family Income in College Selection and Beyond," *Federal Reserve Bank of St. Louis Review*, 105, 10.20955/r.105.198-222.

LOCHNER, L. J., AND A. MONGE-NARANJO (2011): "The Nature of Credit Constraints and Human Capital," *The American Economic Review*, 101, 2487–2529.

LOVENHEIM, M., AND J. SMITH (2023): "Returns to different postsecondary investments: Institution type, academic programs, and credentials," in *Handbook of the Economics of Education* Volume 6: Elsevier, 187–318.

MARTO, R., AND Y. WITTMAN (2024): "Risky College Admissions and the (Mis)allocation of Talent."

PAGE, L. C., AND J. SCOTT-CLAYTON (2016): "Improving college access in the United States: Barriers and policy responses," *Economics of Education Review*, 51, 4–22.

RODERICK, M., V. COCA, AND J. NAGAOKA (2011): "Potholes on the Road to College: High School Effects in Shaping Urban Students' Participation in College Application, Four-year College Enrollment, and College Match," *Sociology of Education*, 84, 178–211, 10.1177/0038040711411280, Publisher: SAGE Publications Inc.

RUPERT, P., AND G. ZANELLA (2015): "Revisiting wage, earnings, and hours profiles," *Journal of Monetary Economics*, 72, 114–130.

Appendix

A Calibration

This section describes the calibration in detail.

Endowments: High school graduates draw a vector (a, p, g, h_1) from a Gaussian copula. To reduce the number of calibrated parameters, we first draw (a, p) from a bivariate Normal distribution with zero means and unit standard deviations. The correlation parameter is calibrated. We then set $g = \beta_{g,a}a + \beta_{g,p}p + \varepsilon_g$ and $h_1 = \beta_{h,a}a + \beta_{h,p}p + \varepsilon_h$, where $\varepsilon_g, \varepsilon_h \sim N(0, 1)$. Finally we scale the four endowments to their respective marginal distributions: standard Normal for a, uniform (percentiles) for g and p, and uniform in $[1, h_{1,max}]$ for h_1. The upper bound $h_{1,max}$ is calibrated.

College admissions: The admissions score z is calculated as $\beta_h h + \beta_g g$, rescaled into percentile values. We normalize $\beta_h = 0.5$ and calibrate β_g.

Calibration algorithm: For each candidate set of parameters, the calibration algorithm calculates the probability of all possible life histories for $10,000$ students. It constructs model counterparts of the target moments and searches for the parameter vector that minimizes a weighted sum of squared deviations between model and data moments.

B Fixed Parameters

How the fixed model parameters are set is described in section Section 3.2. Figure 8 shows the estimated experience-wage profiles. Table 4 through Table 6 show college costs τ, parental transfers z, and student debt k.

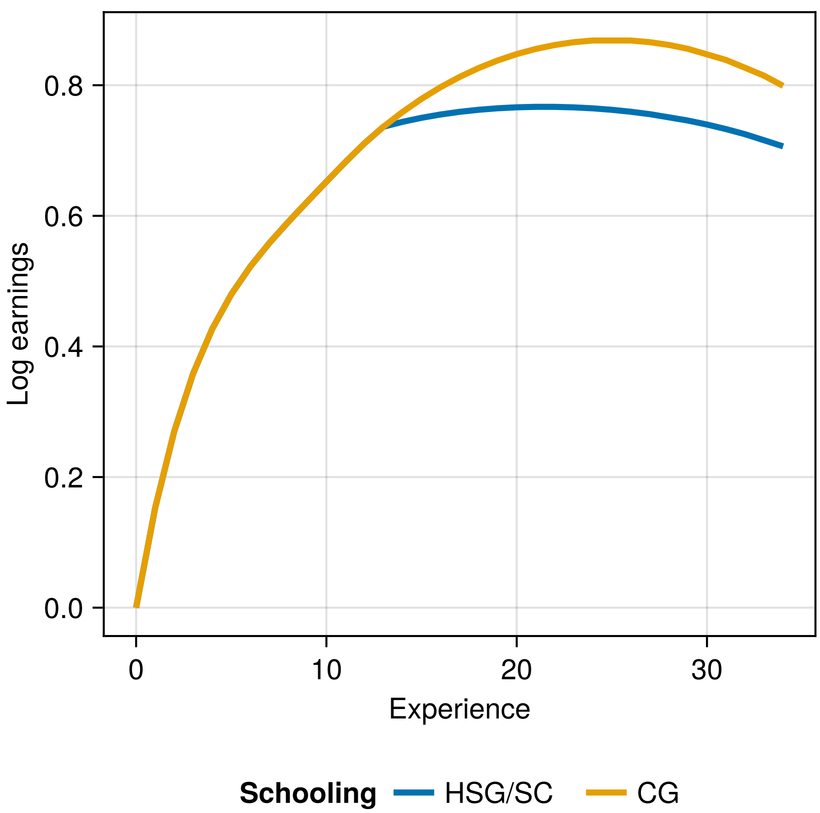
C Calibrated Parameters

Table 7 through Table 9 show the calibrated values for parameters values that have interpretable values. The details are as follows.

Student endowments are drawn from a Gaussian copula. In order to reduce the number of calibrated parameters that govern endowment correlations, we proceed as follows.

1. We draw $(a, p) \sim N(0, \rho_{a,p})$.
2. We set $h_1 = \beta_{h,a}a + \beta_{h,p}p + \varepsilon_h$.
3. We set $g = \beta_{g,a}a + \beta_{g,p}p + \varepsilon_g$.

Figure 8: Experience Profiles



Note: The figure shows the estimated experience profiles for log earnings.

Table 4: College Costs

Regressor	Coefficient (s.e.)
AFQT 2	-261.65
	(615.46)
AFQT 3	39.27
	(573.69)
AFQT 4	-1,198.1
	(629.19)
Quality 2	35.69
	(600.22)
Quality 3	1,679.4
	(546.53)
Quality 4	4,307.6
	(882.17)
Parental income 2	921.85
	(1,005.6)
Parental income 3	1,825.0
	(974.64)
Parental income 4	2,938.7
	(1,023.4)
Constant	-720.65
	(801.01)

Note: The table shows the results of regressing the net cost of college on AFQT quartile dummies, parental quartile dummies, and quality dummies. The regression implies negative net costs for some students with low incomes but high test scores. Such cases, where financial aid covers more than the total cost of college, are observed in the data.

Table 5: Parental Transfers

	$Y1$	$Y2$	$Y3$	$Y4$
Quality 1	626.6	1,032.1	2,442.7	2,796.4
Quality 2	2,079.1	2,237.4	4,186.2	5,274.6
Quality 3	1,700.3	3,913.6	5,289.6	9,998.7
Quality 4	4,171.0	4,896.5	11,095.6	17,218.7

Note: The table shows mean parental transfers for each combination of family income quartile and college quality.

Table 6: Student Debt

Year	**1**	**2**	**3**	**4**
Quality 1	374.0	794.2	n/a	n/a
Quality 2	2,569.5	5,671.5	7,756.1	10,773.1
Quality 3	2,070.5	4,515.3	7,159.0	10,064.8
Quality 4	2,717.5	5,395.5	7,574.9	11,121.2

Note: The table shows mean student debt at the end of each year in college for each quality.

Table 7: Preference parameters

Symbol	Description	Value
$\mathcal{U}_e$	Fixed utility at work; by education	3.00, 2.63, 3.50
$\mathcal{U}_{2y}$	Utility from attending 2 year college	6.97
$\Delta\mathcal{U}$	Range of idiosyncratic college preferences	5.14

4. Admissions scores are given by $z = 0.5h_1 + \beta_{z,g}g + \varepsilon_z$.
5. Finally, we rescale the endowments to have the desired marginal distributions: $a \sim N(0,1)$, $p \sim U[0,1]$, $h_1 \sim U[1, 1 + \Delta h_1]$, , and $g \sim U[0,1]$.

   All of the error terms $(\varepsilon_h, \varepsilon_g, \varepsilon_z)$ are drawn from independent, standard Normal distributions.

Some of the college related parameters are not easily interpreted. Their implications are shown in Figure 9a through Figure 1.

1. Students' graduation and dropout probabilities ($\Pr_g$ and $\Pr_d$) are shown in Figure 9a and Figure 9b, respectively. For a given student, attending a higher quality college increases the likelihood of graduating, but reduces the likelihood of dropping out. Taken together, the probability of eventually graduating from college rises substantially with college quality, especially for high-ability students.
2. For each college, learning productivity $\mathcal{A}(q, a)$ is a function of student ability (see equation (6)). The implied productivities are shown in Figure 1.

D Model Fit

This section shows all target moments used in the calibration, except for those already displayed in Section 3.4.

D.1 College Entry Patterns

The target moments that characterize college entry patterns are:

1. The fraction of high school graduates in each AFQT or parental income quartile who enter any college (Figure 10 and Figure 11).
2. The fraction of college entrants in each AFQT and parental income quartile who choose each college (Figure 3 through Figure 4 and Figure 12 through Figure 15).

Table 8: Endowment Parameters

Symbol	Description	Value
$\rho_{a,p}$	Correlation (a,p)	0.340
$\beta_{h,a}$	Weight on ability when drawing h_1	1.30
$\beta_{h,p}$	Weight on parental when drawing h_1	0.479
Δh_1	Range of h endowments	0.120
$\beta_{g,a}$	Weight on ability when drawing g	3.71
$\beta_{g,p}$	Weight on parental when drawing g	0.0298
$\beta_{z,g}$	Weight on g in admissions score	0.119
π	Prob of observing true quality	0.338, 0.406, 0.470, 0.526

Note: The probability of observing the true college quality varies with parental income quartile.

Table 9: Financial Parameters

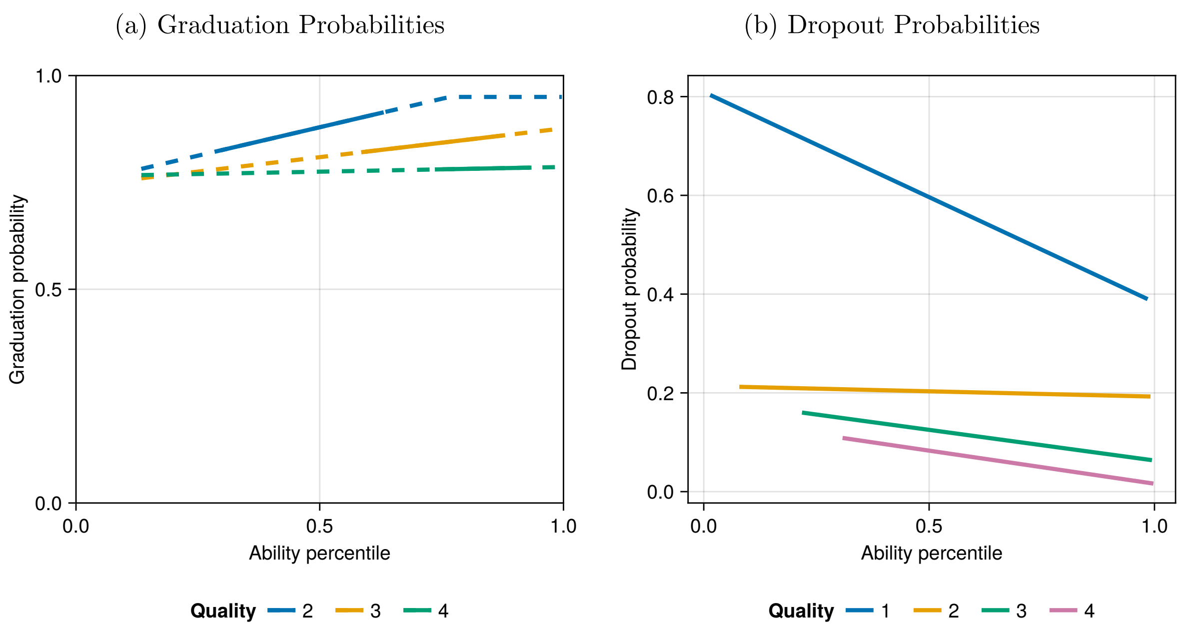
Symbol	Description	Value
τ_{4y}	Cost of attending four year college	4.11
w_{HSG}	Log wage HSG	2.39
Δw	College wage premium	0.0692

Note: The college wage premium is the log difference between w_{CG} and w_{HSG}.

Table 10: College Related Parameters

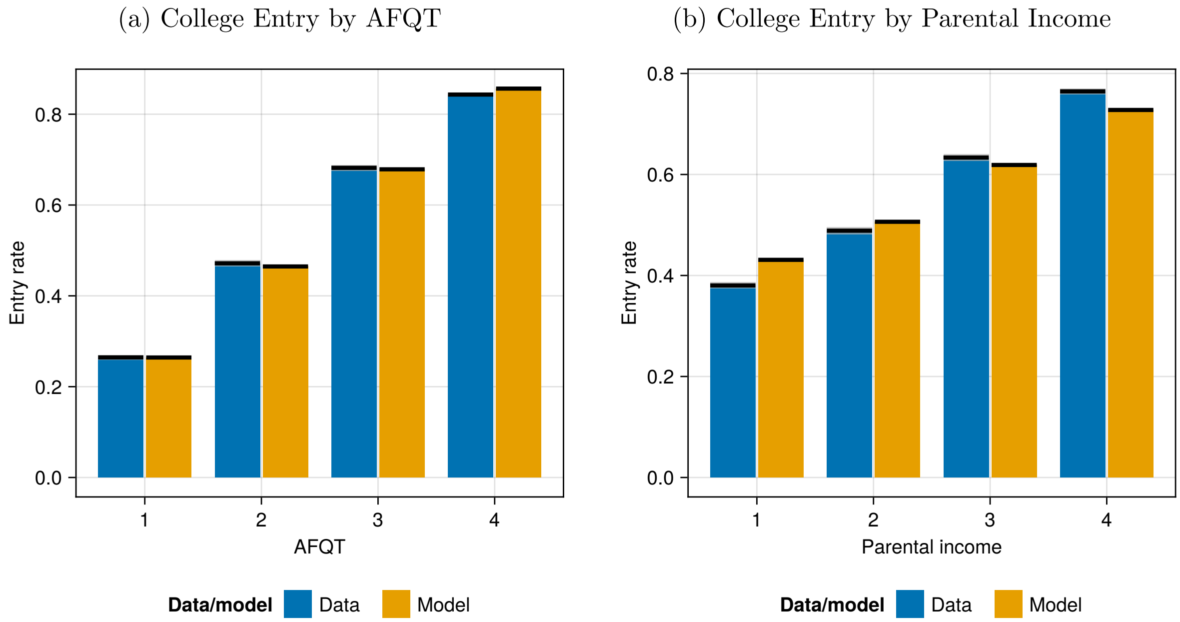
Symbol	Description	Value
A_q	College productivities	-2.82, -2.02, -1.95, -1.95
ϕ_q	Ability scale	0.000, 0.000, 0.164, 0.164
ϕ	Ability squared scale	0.141
ζ	Exponent on h	0.054
$\gamma_{1,q}$	Graduation probability intercepts	0.000, 0.746, 0.741, 0.764
$\gamma_{2,q}$	Graduation probability slopes	0.000, 0.266, 0.136, 0.022
$\gamma_{4,q}$	Dropout probability intercepts	0.809, 0.214, 0.187, 0.151
$\gamma_{5,q}$	Dropout probability slopes	-0.426, -0.022, -0.124, -0.134

Figure 9: Graduation and Dropout Probabilities



Note: The figure shows the graduation and dropout probabilities for students of different ability levels. The solid sections in the graduation probability plot represent inter-quartile ranges of students enrolled in each college.

Figure 10: College Entry by AFQT or Parental Income



3. Mean AFQT percentiles of freshmen in each college (Figure 16).
4. Total freshman enrollment by college quality (Figure 17).

D.2 College Dropout and Graduation

Target moments that relate to college dropout and graduation patterns are:

1. The fraction of college entrants that graduate within six years by college quality, AFQT quartile, and parental income quartile (Figure 18 through Figure 20).
2. The fraction of college entrants dropping out by the end of the second year by AFQT and college quality (Figure 21).
3. The fraction of college entrants dropping out out at the end of each year for each college quality (Figure 22).
4. The average number of years students spend in college before either dropping out or graduating (Figure 23).

D.3 Other Target Moments

Target moments that characterize worker earnings are:

Figure 11: College Entry by AFQT and Parental Income

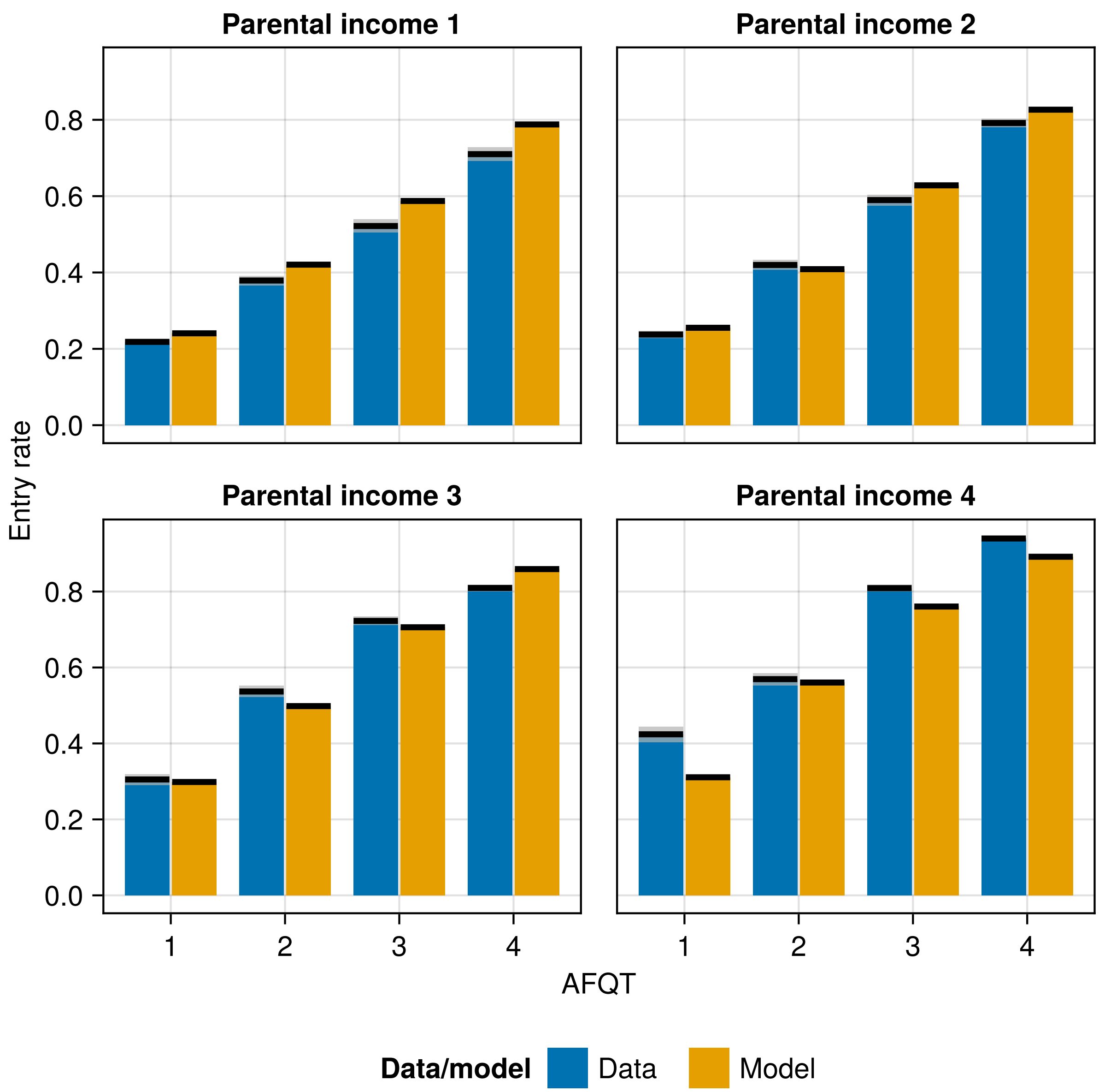


Figure 12: College Quality Choice for Parental Income Quartile 1

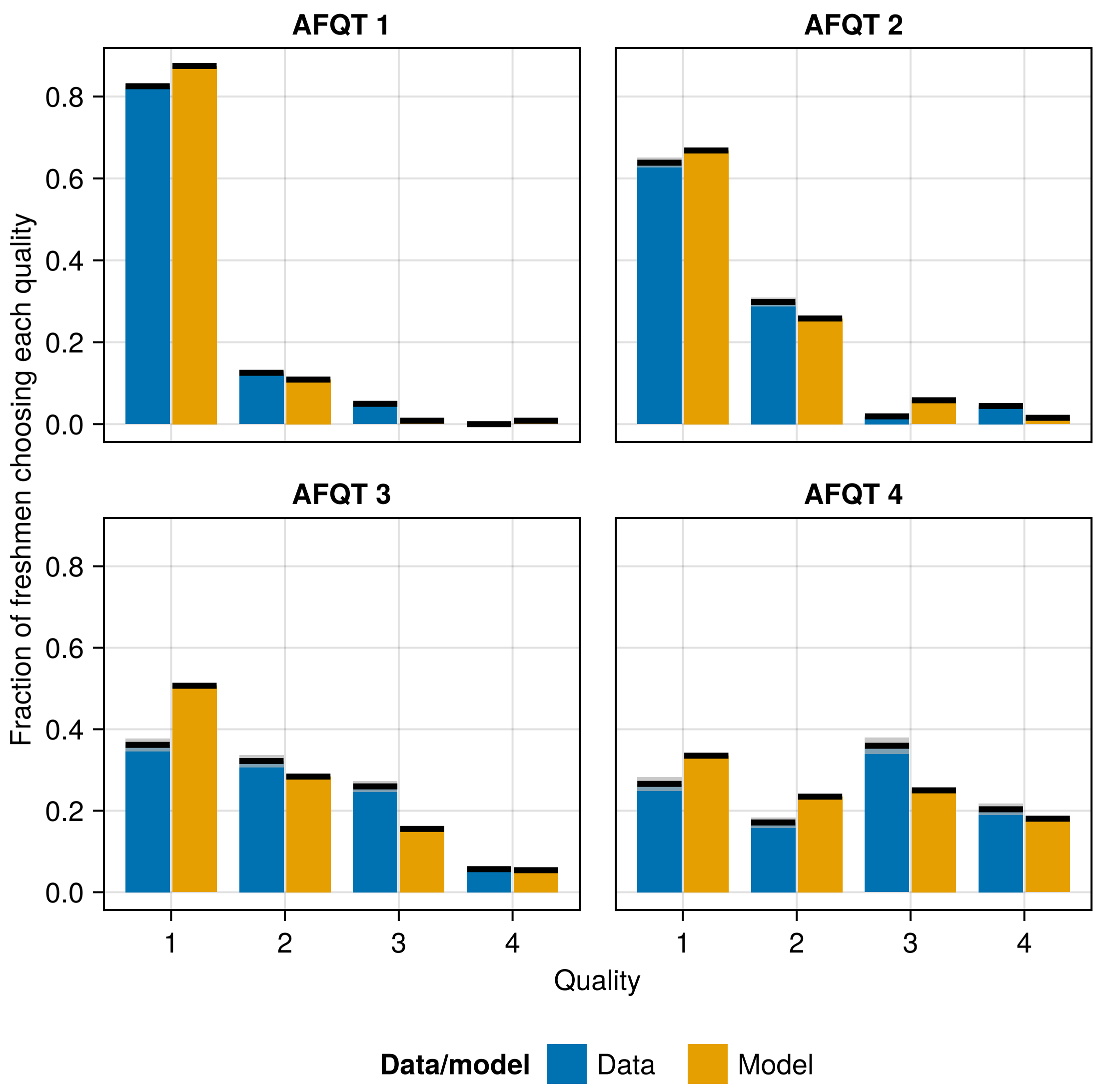


Figure 13: College Quality Choice for Parental Income Quartile 2

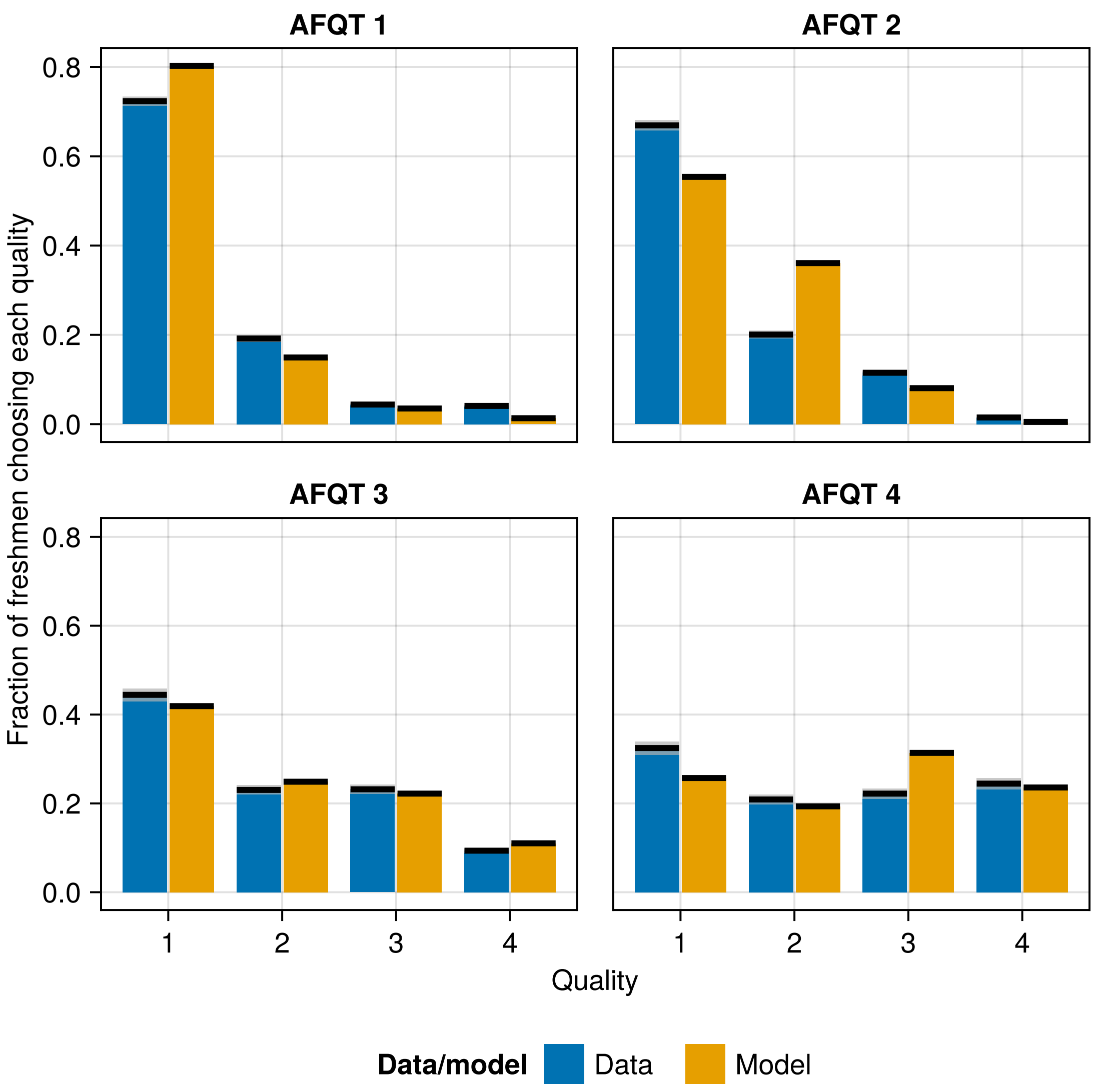


Figure 14: College Quality Choice for Parental Income Quartile 3

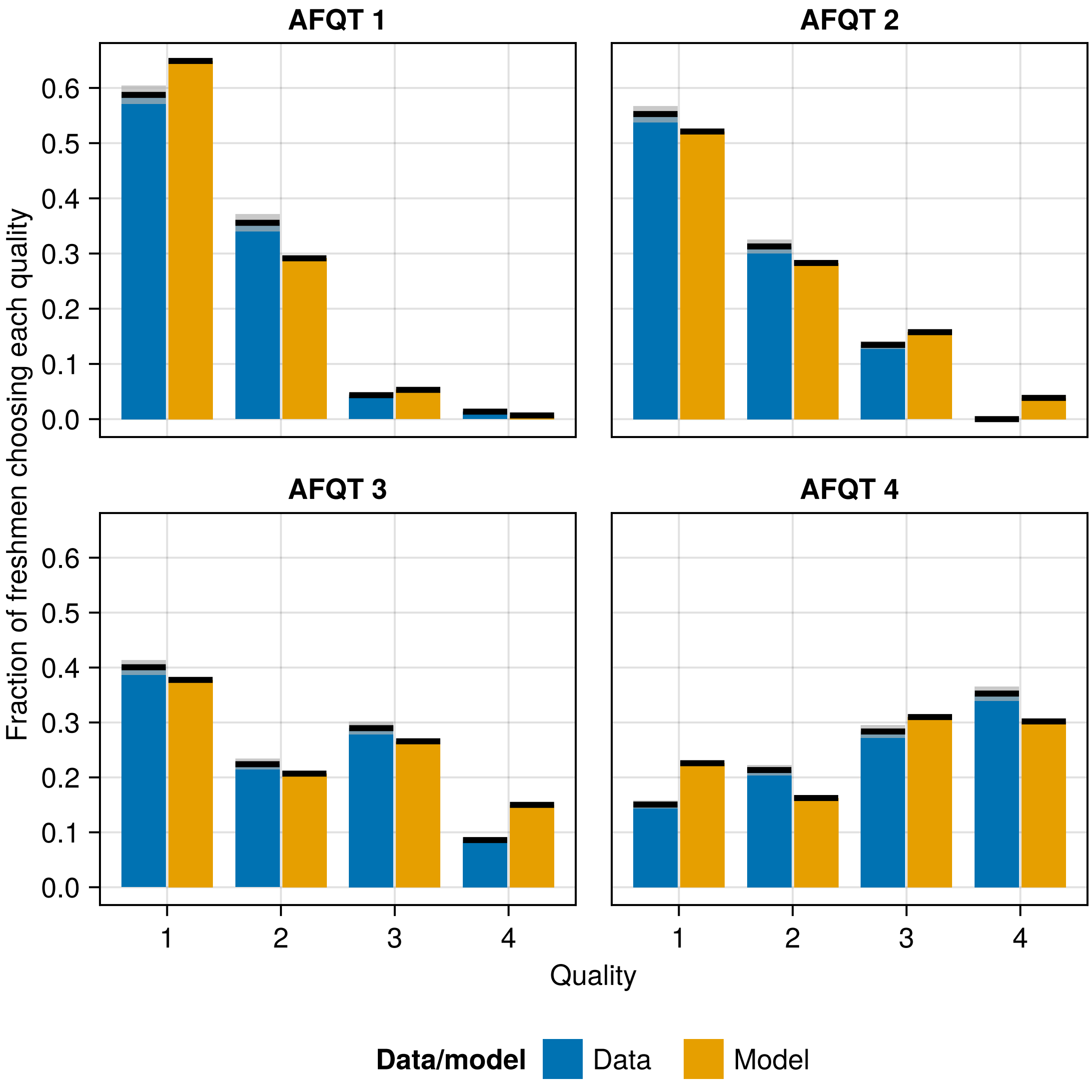


Figure 15: College Quality Choice for Parental Income Quartile 4

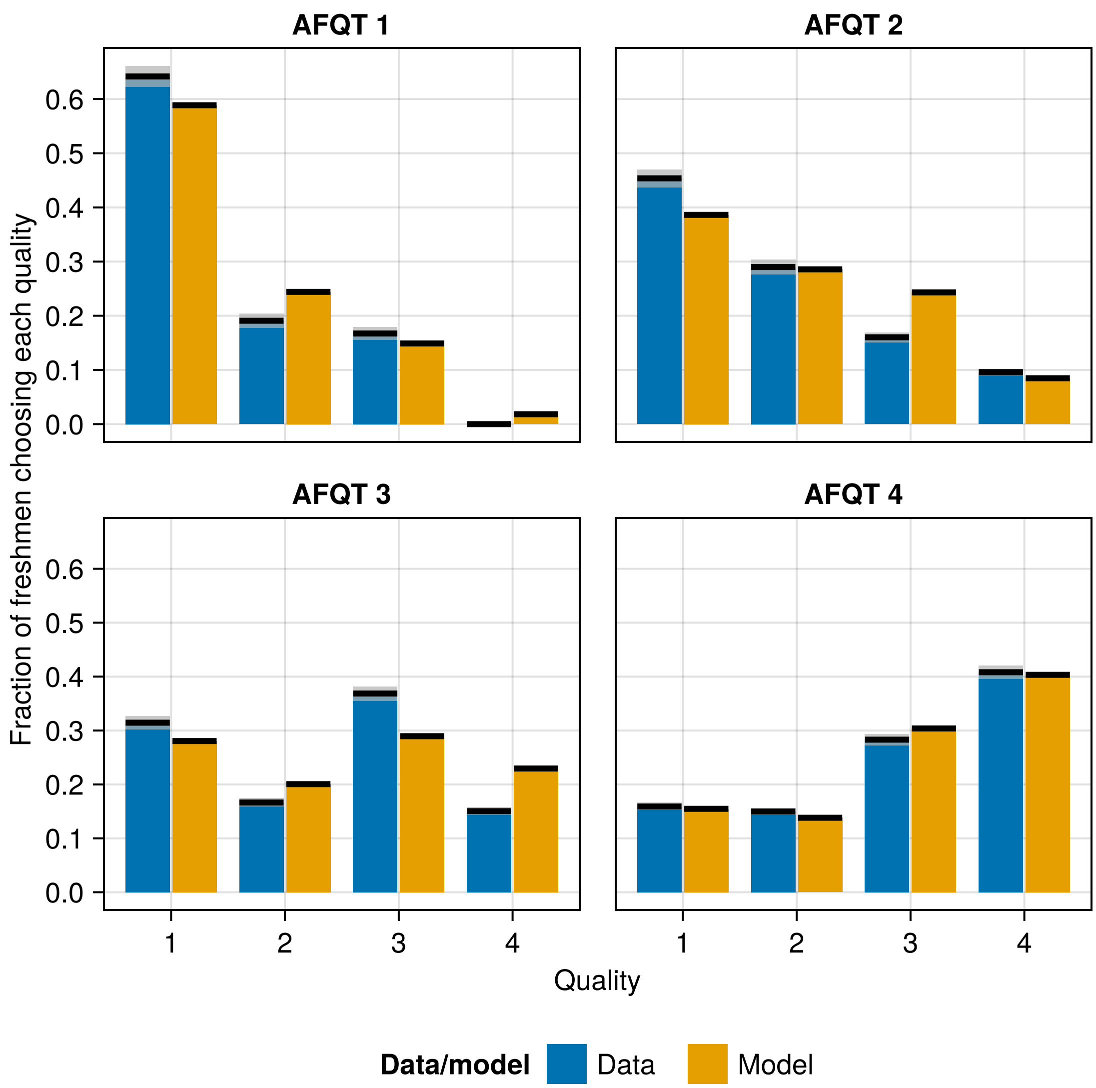


Figure 16: Mean AFQT Percentiles

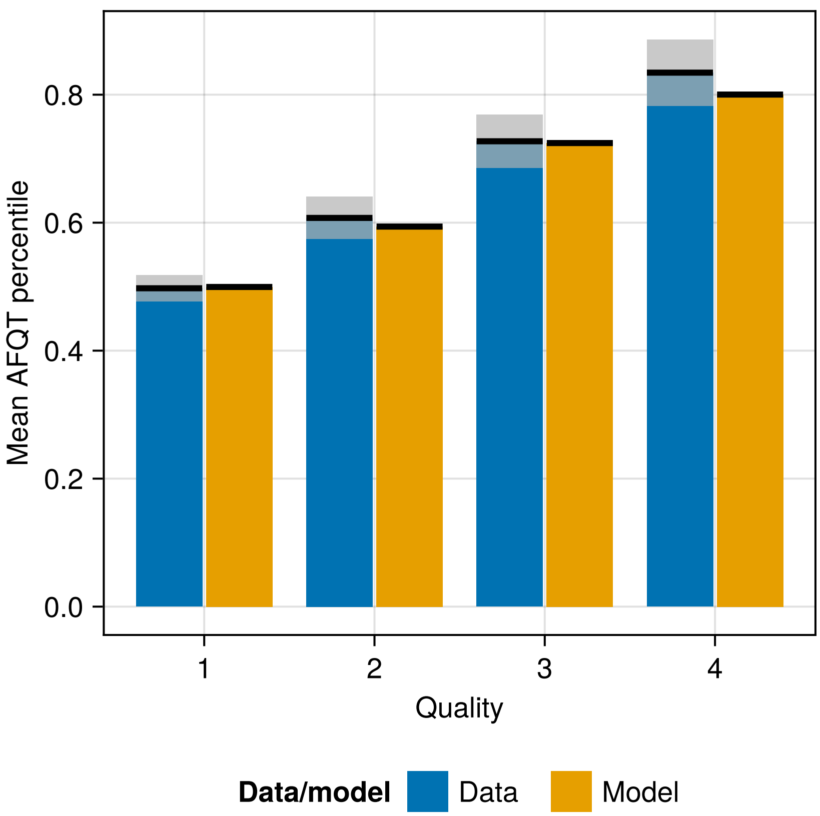


Figure 17: Enrollment by College Quality

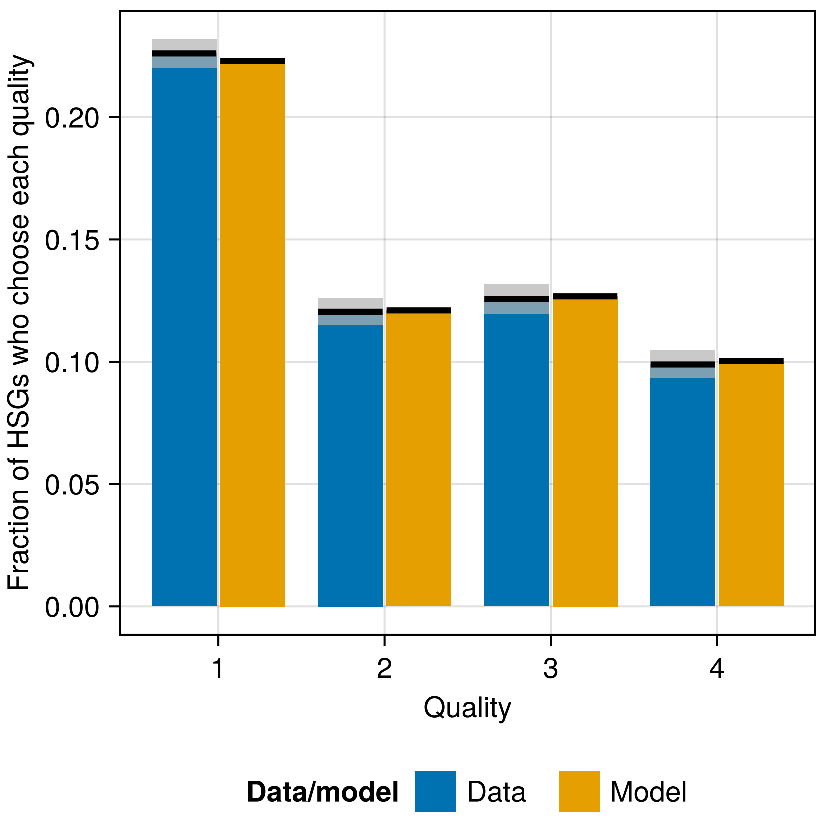
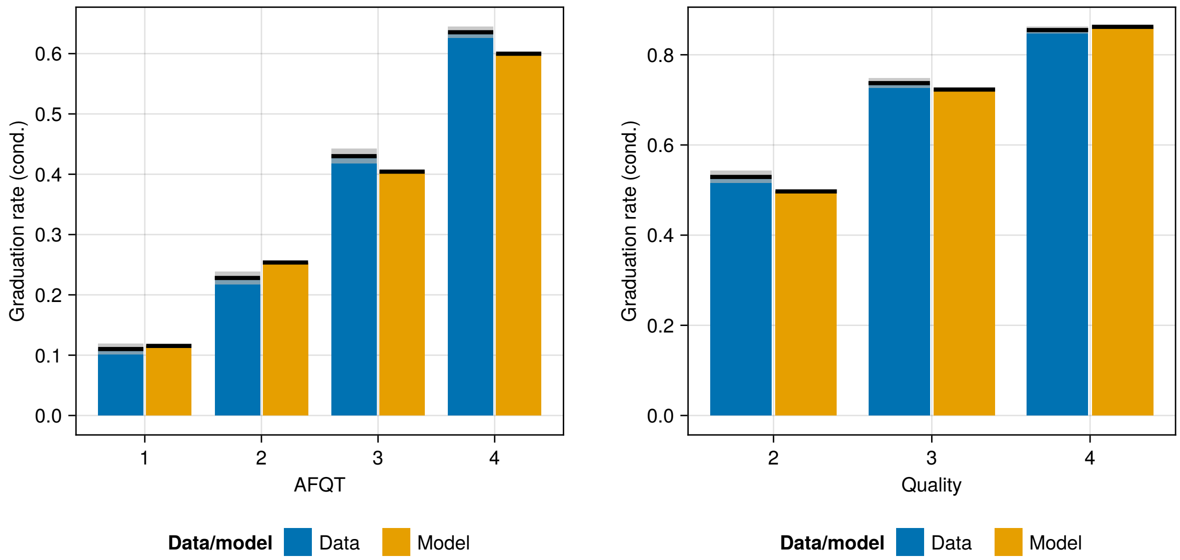


Figure 18: Graduation Rates

(a) By AFQT

(b) By College Quality



1. The coefficients of a regression of log earnings (net of experience effects) on AFQT and education dummies (Table 11). Even controlling for AFQT scores, college graduates earn far more than dropouts.
2. Mean log earnings fixed effects by education, AFQT, and college quality (Figure 24 through Figure 26).

Scalar target moments are shown in Table 12. The last two rows show the quasi-experimental moments described in Section 3.

1. The "tuition increase" entry shows the change in college enrollment due to a \$5,000 increase in tuition. A random subset of 40 percent of high school graduates receive the treatment. The target moment is based on Dynarski et al. (2022b).
2. The "full information" entry shows the change in college enrollment for lower-income, high AFQT students who are given full information ($\pi = 1$). The target moment is based on Hoxby et al. (2013).

In both cases, the enrollment changes are calculated as the difference between the mean enrollment change of the treated and the untreated students.

Figure 19: Graduation Rates

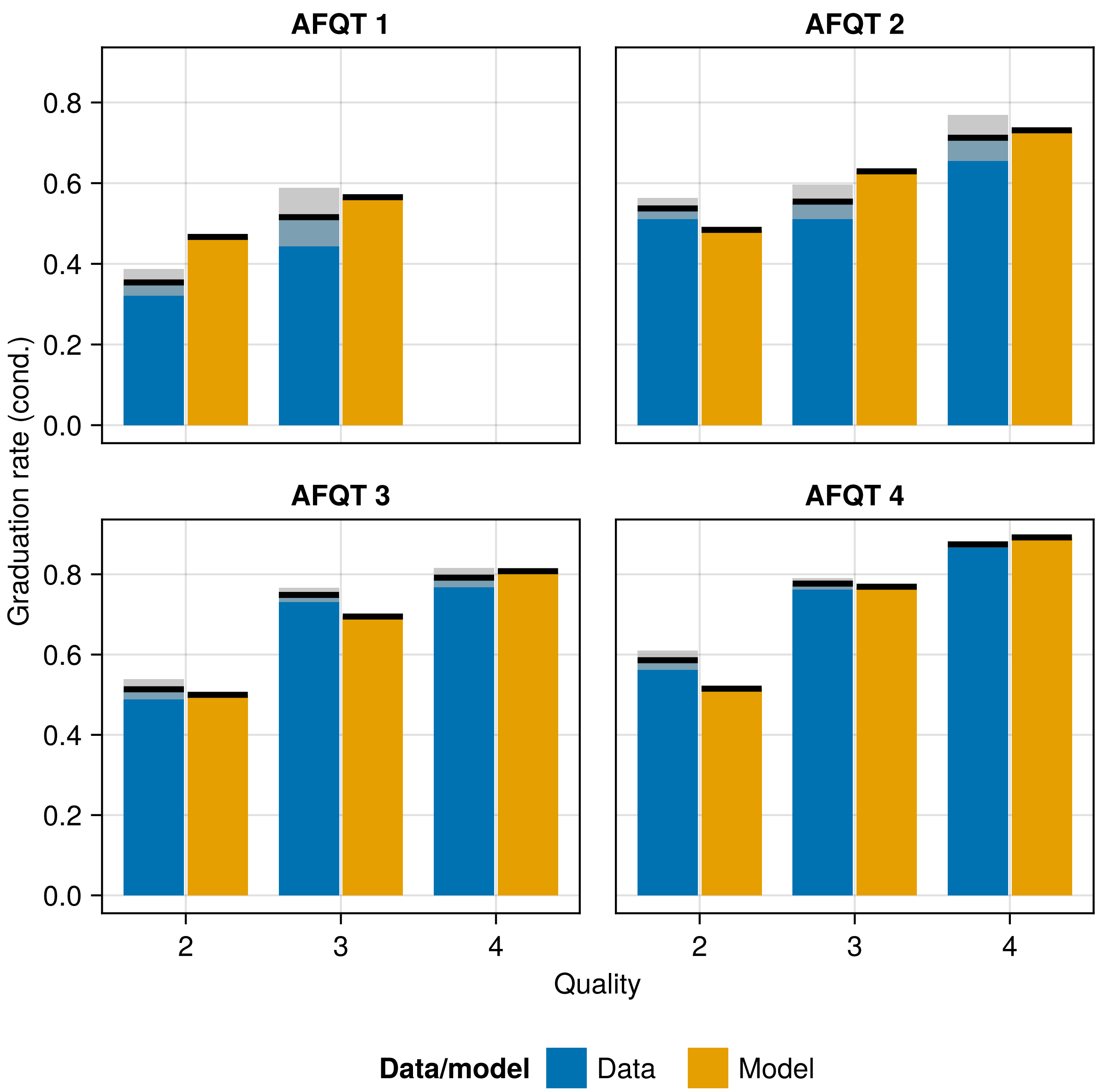


Figure 20: Graduation Rates by Quality and Parental Income

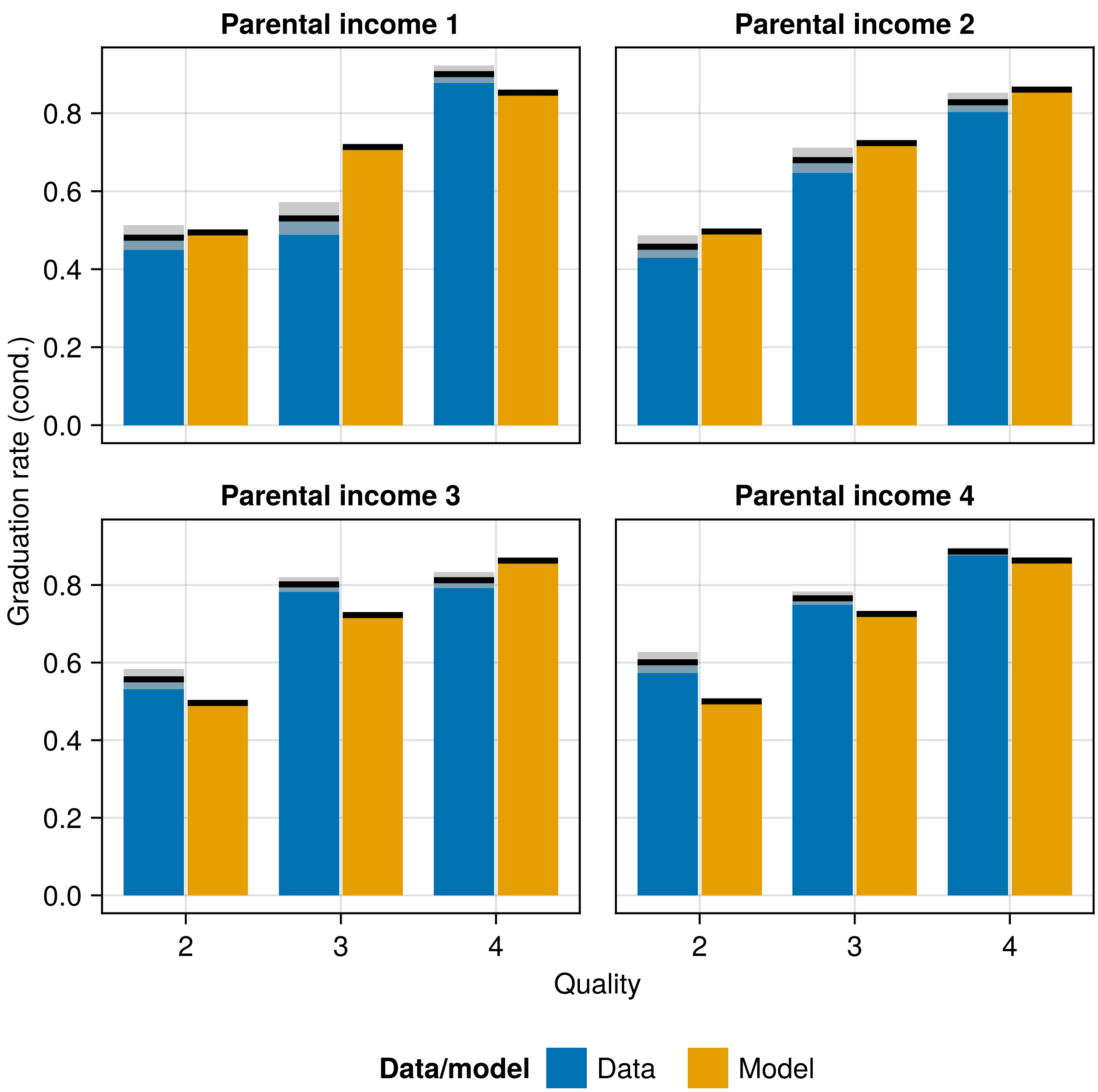


Figure 21: Cumulative Dropout Rates at End of Year 2

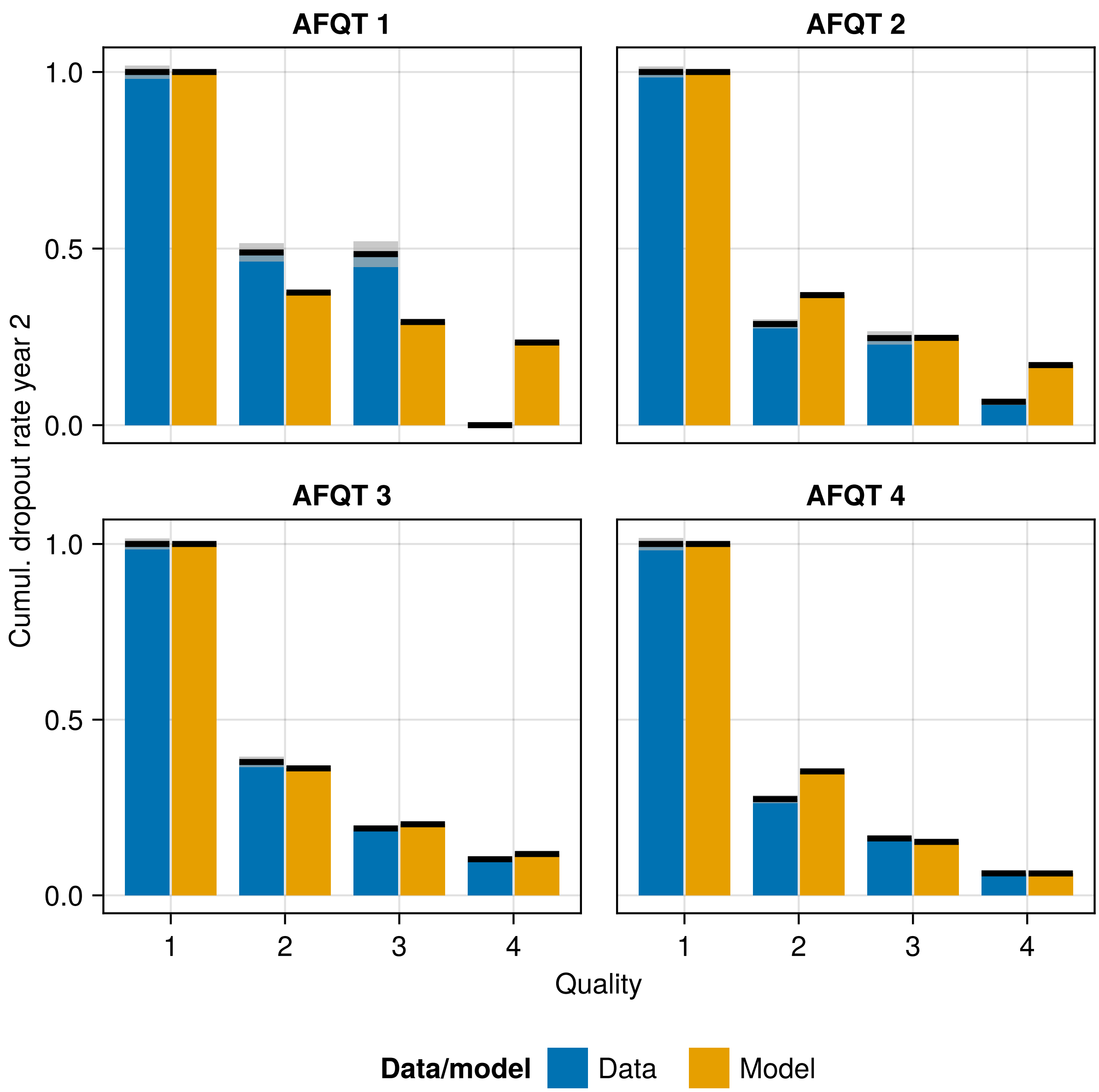


Figure 22: Fraction of Entrants that Drop Out by Year

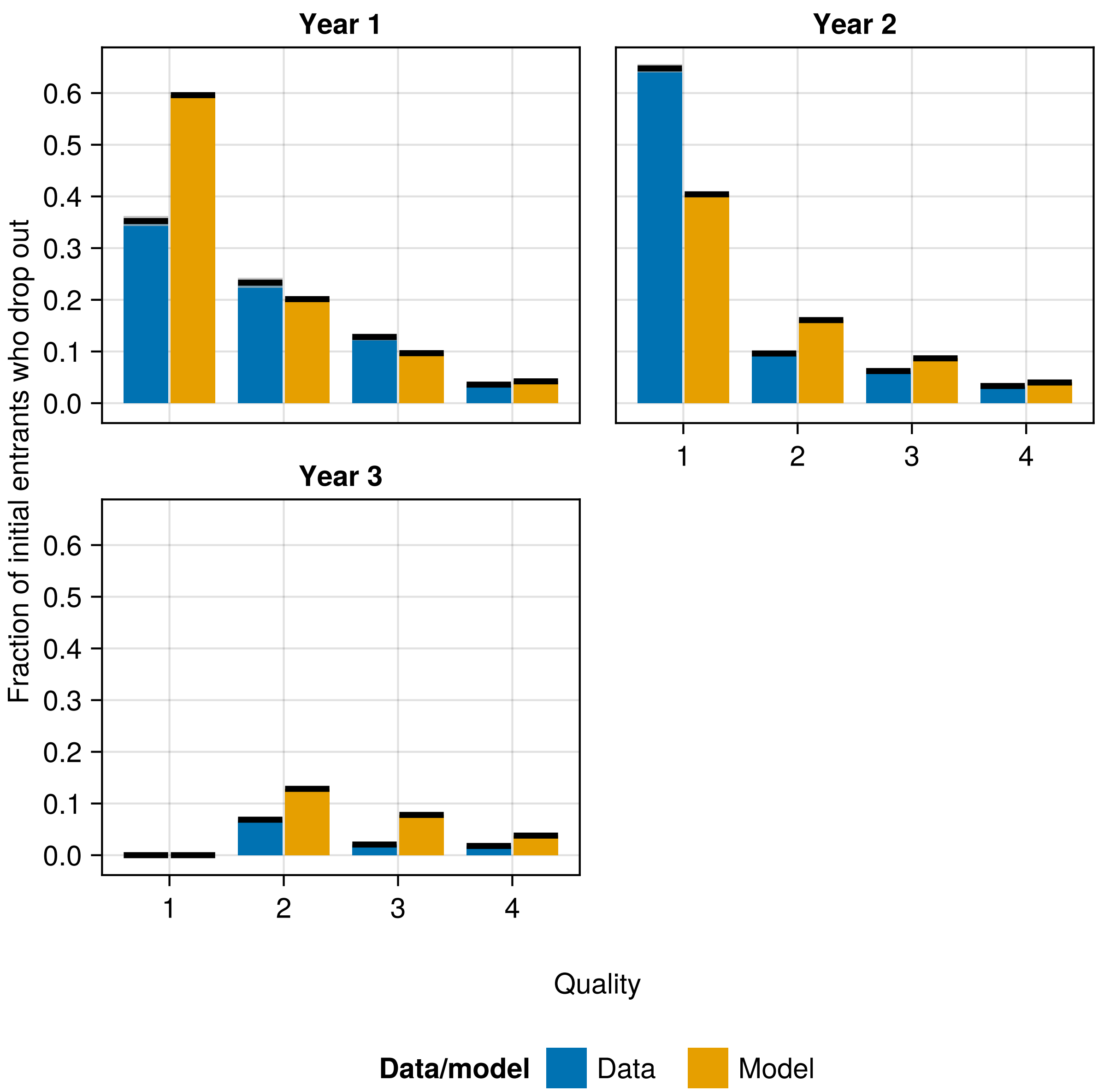
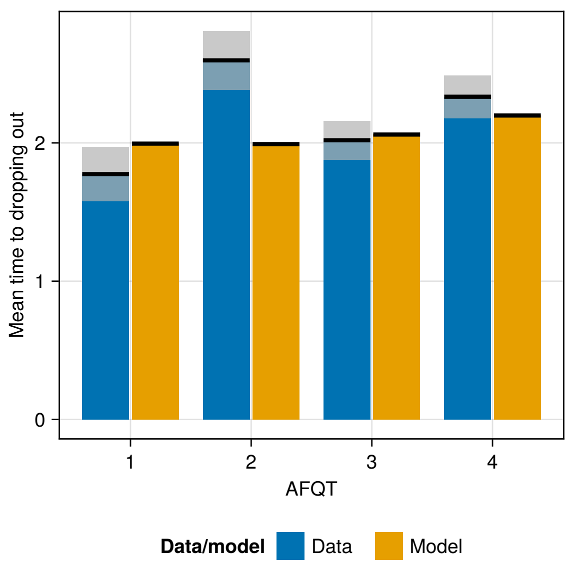
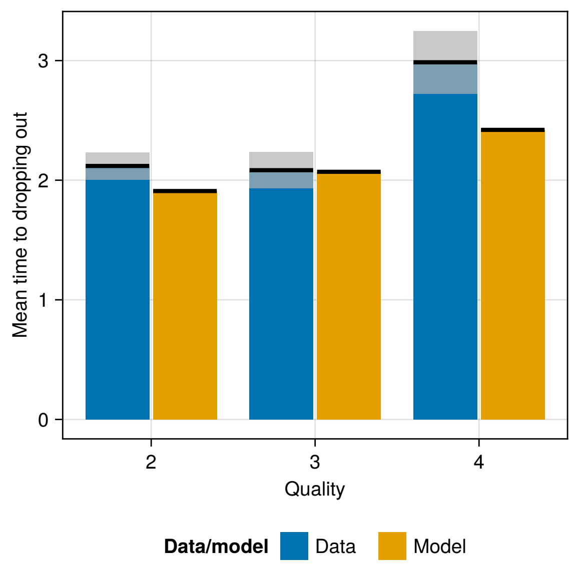


Figure 23: Mean Time to Dropout and Graduation

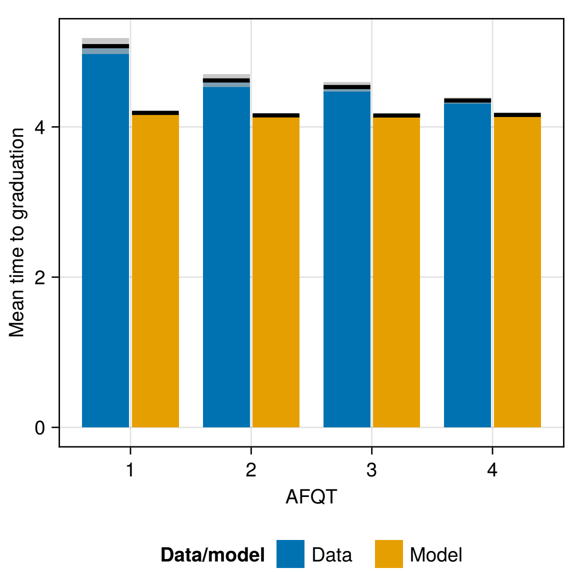
(a) Mean Time to Dropping Out



(b) Mean Time to Dropping Out



(c) Mean Time to Graduation



(d) Mean Time to Graduation

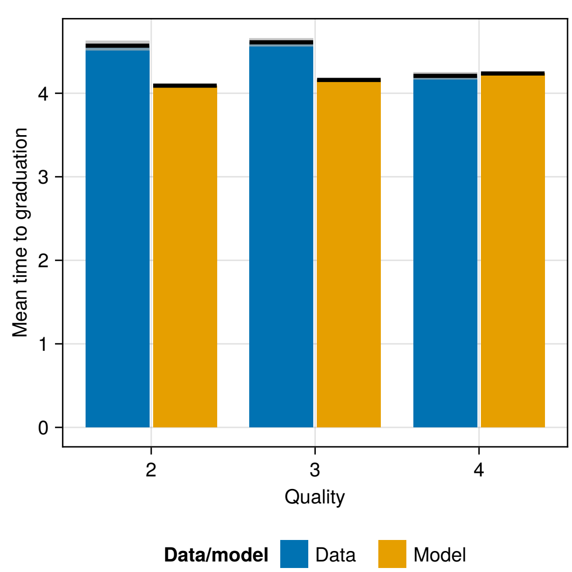


Table 11: Earnings Regressions. All Workers.

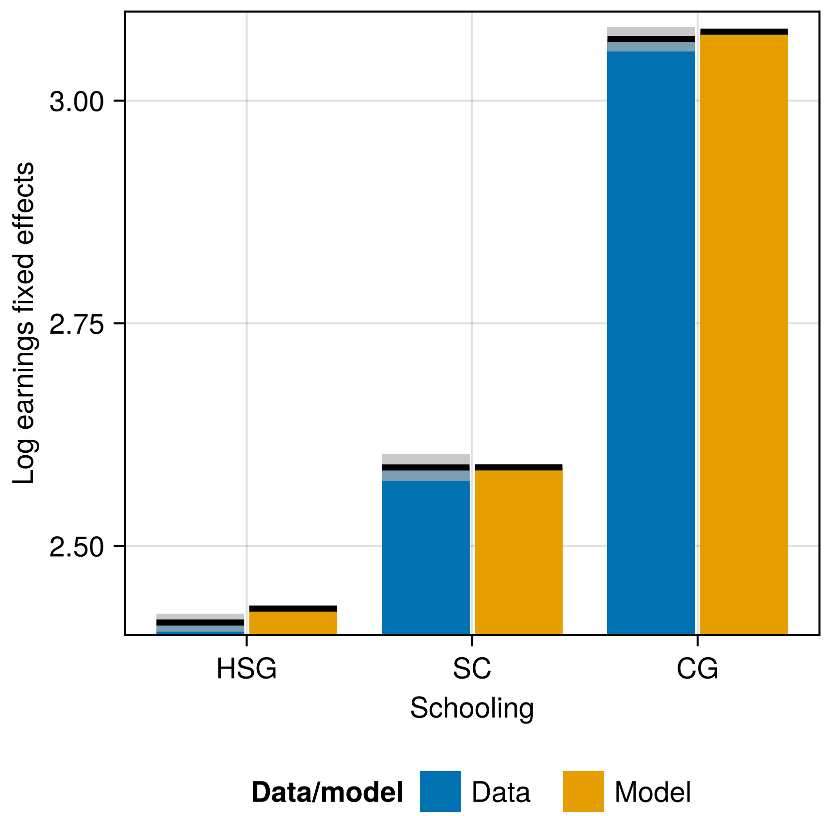
Regressor	**Data**	**Model**
AFQT 2	0.102	0.0220
	(0.0198)	
AFQT 3	0.121	0.0441
	(0.0208)	
AFQT 4	0.187	0.128
	(0.0232)	
SC	0.147	0.124
	(0.0193)	
CG	0.583	0.573
	(0.0182)	
Constant	2.34	2.40
	(0.0145)	

Table 12: Scalar Target Moments

Description	**Data**	**Model**
Graduation rate (cond.)	0.44 (0.01)	0.42
Entry rate	0.57 (0.01)	0.57
Tuition increase (Dynarski et al. 2022)	17.5	17.5
Information intervention (Hoxby and Turner 2013)	5.3	5.32

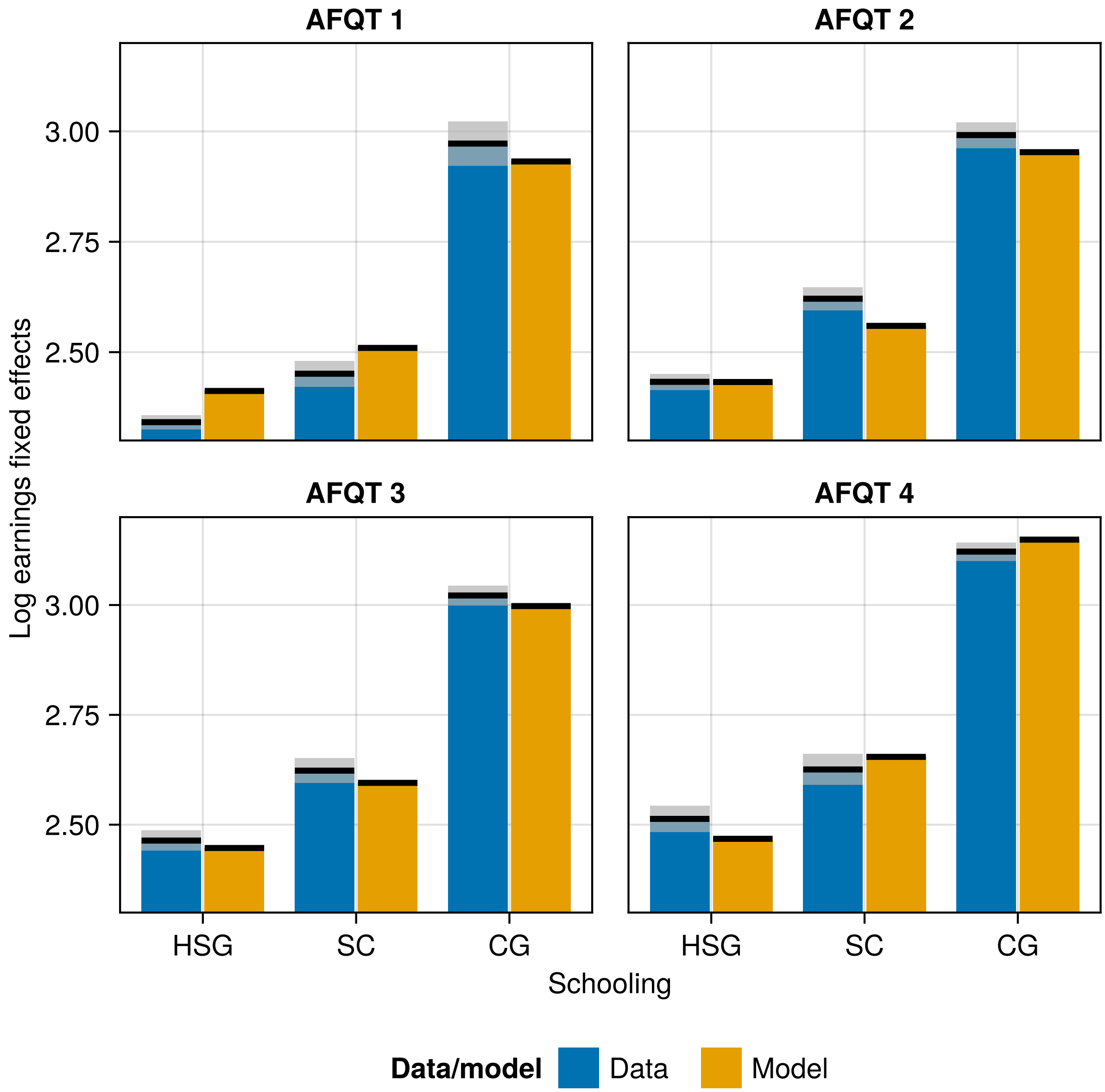
Note: Standard errors for data moments are shown in parentheses where applicable.

Figure 24: Wage Fixed Effects by Schooling



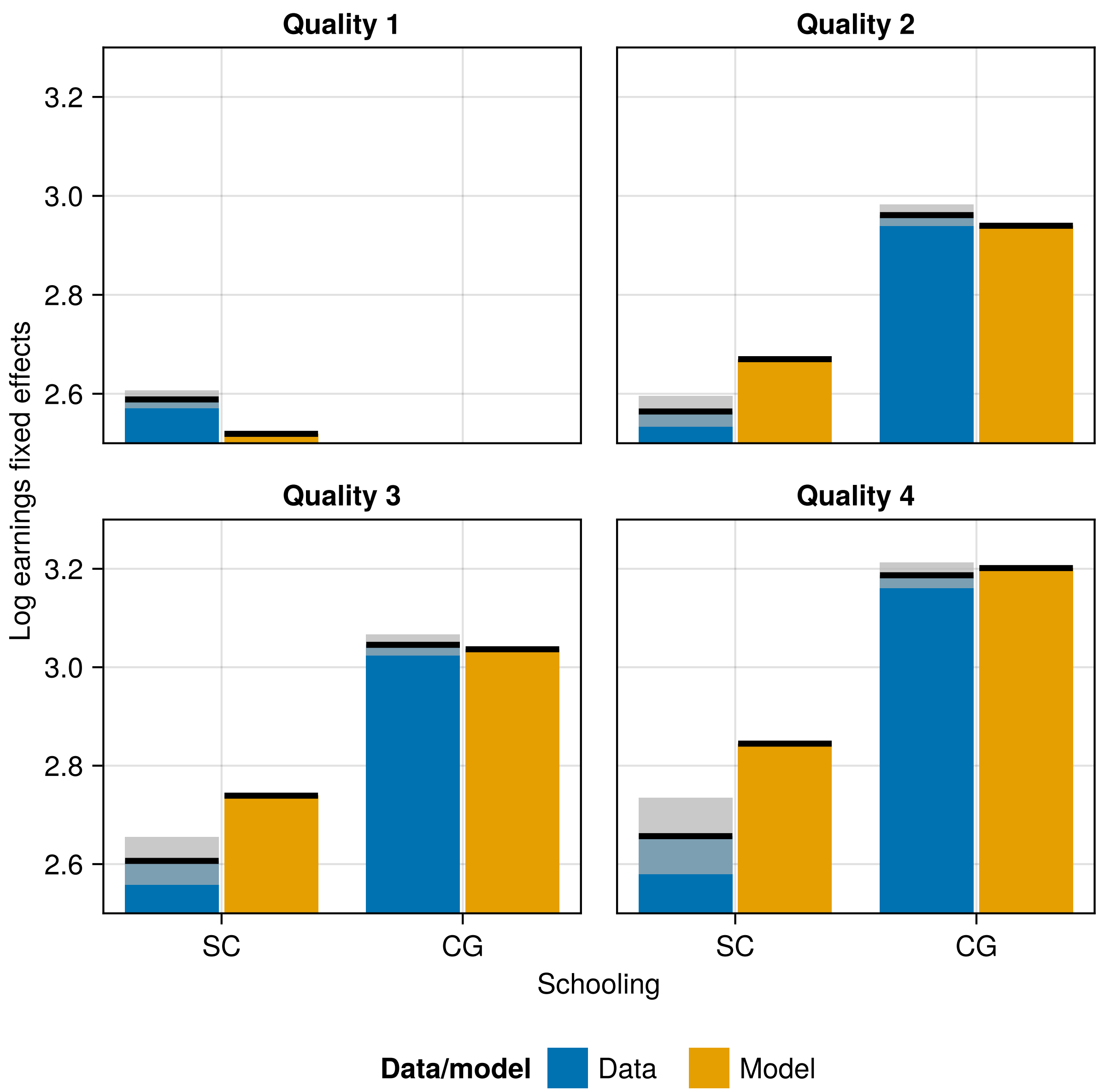
Note: In the data, fixed effects are estimated by regressing log earnings on a quartic polynomial in potential experience. The regressions use panel data and are estimated separately by eduction group. See Leukhina (2023) for details. In the model, individual log earnings are given by log earnings in the first year of work plus an experience profile that is common to all workers of a given education group. The fixed effect is then simply the level of initial earnings.

Figure 25: Wage Fixed Effects by Schooling and AFQT



Note: See Figure Figure 24 for a definition of earnings fixed effects.

Figure 26: Earnings Fixed Effects by Schooling and College



Note: See Figure Figure 24 for a definition of earnings fixed effects.

D.4 Instructional Expenditure Categories

Our college expenditure data are derived from the Integrated Postsecondary Education Data System (IPEDS), a system of surveys conducted each year by the National Center for Education Statistics (NCES). Any institution that participates in a federal student financial aid program is required to report their data to the NCES. Thus, IPEDS is a dataset of all colleges and universities in the U.S.

IPEDS contains financial data of both public and private institutions. In general, private institutions use accounting standards established by the Financial Accounting Standards Board (FASB), whereas public institutions follow the accounting standards of the Governmental Accounting Standards Board (GASB). We utilize finance data of two- and four-year colleges for years 2003-2005. Expenditures are inflation-adjusted to 2000 dollars using the CPI.

We derive our measure of total instruction-related expenditures by summing over the following categories of expenses: instruction, academic support, institutional support, student services, and auxiliary expenses. The aggregate expenditure measure is divided by the school's 12-month full time equivalent enrollment (FTE) to obtain expenditure per student.

We generate the mean college expenditure for each college type using the 12-month FTE as weights. Schools missing any of the expenditure categories are not considered when calculating these means.

Expenditure categories are described as follows:

Instruction expenses - A functional expense category that includes expenses of the colleges, schools, departments, and other instructional divisions of the institution and expenses for departmental research and public service that are not separately budgeted. Includes general academic instruction, occupational and vocational instruction, community education, preparatory and adult basic education, and regular, special, and extension sessions. Also includes expenses for both credit and non-credit activities. Excludes expenses for academic administration where the primary function is administration (e.g., academic deans). Information technology expenses related to instructional activities if the institution separately budgets and expenses information technology resources are included (otherwise these expenses are included in academic support). FASB institutions include actual or allocated costs for operation and maintenance of plant, interest, and depreciation. GASB institutions do not include operation and maintenance of plant or interest, but may, as an option, distribute depreciation expense.

Student services expenses - A functional expense category that includes expenses for admissions, registrar activities, and activities whose primary purpose is to contribute to students emotional and physical well - being and to their intellectual, cultural, and social development outside the context of the formal instructional program. Examples include student activities, cultural events, student newspapers, intramural athletics, student organizations, supplemental instruction outside the normal administration, and student records. Intercollegiate athletics and student health services may also be included except when operated as self - supporting auxiliary enterprises. Also may include information technology expenses related to student service activities if the institution separately budgets and expenses information technology resources (otherwise these expenses are included in institutional support.) FASB institutions include actual or allocated costs for operation and maintenance of plant, interest, and depreciation. GASB institutions do not include operation and maintenance of plant or interest but may, as an option, distribute depreciation expense.

Academic support expenses - A functional expense category that includes expenses of activities and services that support the institution's primary missions of instruction, research, and public service. It includes the retention, preservation, and display of educational materials (for example, libraries, museums, and galleries); organized activities that provide support services to the academic functions of the institution (such as a demonstration school associated with a college of education or veterinary and dental clinics if their primary purpose is to support the instructional program); media such as audiovisual services; academic administration (including academic deans but not department chairpersons); and formally organized and separately budgeted academic personnel development and course and curriculum development expenses. Also included are information technology expenses related to academic support activities; if an institution does not separately budget and expense information technology resources, the costs associated with the three primary programs will be applied to this function and the remainder to institutional support. Under FASB standards this includes actual or allocated costs for operation and maintenance of plant, interest, and depreciation. Under GASB standards this does not include operation and maintenance of plant or interest but may include depreciation expense.

Institutional support expenses - A functional expense category that includes expenses for the day-to-day operational support of the institution. Includes expenses for general administrative services, central executive-level activities concerned with management and long range planning, legal and fiscal operations, space management, employee personnel and records, logistical services such as purchasing and printing, and public relations and development. Also includes information technology expenses related to institutional support activities. If an institution does not separately budget and expense information technology resources, the costs associated with student services and operation and maintenance of plant will also be applied to this function. FASB institutions include actual or allocated costs for operation and maintenance of plant, interest and depreciation. GASB institutions do not include operation and maintenance of plant or interest, but may, as an option, distribute depreciation expense.

Auxiliary enterprise - The sum of all operating expenses associated with essentially self-supporting operations of the institution that exist to furnish a service to students, faculty, or staff, and that charge a fee that is directly related to, although not necessarily equal to, the cost of the service. Examples are residence halls, food services, student health services, intercollegiate athletics (only if essentially self-supporting), college unions, college stores, faculty and staff parking, and faculty housing.