The Review of Economic Studies Advance Access published May 15, 2015

College-Major Choice to College-Then-Major Choice*

Paola Bordon Chao Fu[†]

Ministry of Education of Chile *University of Wisconsin-Madison*

March, 2015

Abstract

Many countries use college-major-specific admissions policies that require a student to choose a college-major pair jointly. Given the potential of student-major mismatches, we explore the equilibrium effects of postponing student choice of major. We develop a sorting equilibrium model under the college-major-specific admissions regime, allowing for match uncertainty and peer effects. We estimate the model using Chilean data. We introduce the counterfactual regime as a Stackelberg game in which a social planner chooses college-specific admissions policies and students make enrollment decisions, learn about their fits to various majors before choosing one. Our estimates indicate that switching from the baseline to the counterfactual regime leads to a 1% increase in average student welfare and that it is more likely to benefit female, low-income and/or low-ability students.

Keywords: College-major choice, major-specific ability, uncertainty, peer effects, equilibrium, admissions systems, cross-system comparison.

*We thank Fumihiko Suga and Yuseob Lee for excellent research assistance. We thank the editor and three anonymous referees for their suggestions. We benefit from discussions with Joe Altonji, Peter Arcidiacono, Steven Durlauf, Hanming Fang, Jim Heckman, Joe Hotz, Mike Keane, John Kennan, Rasmus Lentz, Fei Li, Costas Meghir, Robert Miller, Antonio Penta, John Rust, Xiaoxia Shi, Alan Sorensen, Chris Taber, Xi Weng, Matt Wiswall and Ken Wolpin, as well as comments from workshop participants at the Cowles Summer Conference 2012, Structural Estimation of Behavioral Models Conference, S&M Workshop at Chicago Fed, Econometric Society summer meeting 2012, Duke, IRP-UW and CDE-UW. All errors are ours.

[†]Corresponding author. Department of Economics, University of Wisconsin-Madison. 1180 Observatory Dr. Madison, WI 53706. Email: cfu@ssc.wisc.edu

1 Introduction

In countries such as Canada and the U.S., students are admitted to colleges without declaring their majors until later in their college life.[1] Peer students in the same classes during early college years may end up choosing very different majors. In contrast, many (if not most) countries use college-major-specific admissions rules. A student is admitted to a specific college-major pair and attends classes with peers (mostly) from her own major. We label the first system where students choose majors after enrollment as Sys.S (for sequential), and the second system where students have to make a joint college-major choice as Sys.J (for joint).

Which system is better for the same population of students? This is a natural and policy-relevant question, yet one without a simple answer. To the extent that college education is aimed at providing a society with specialized personnel, Sys.J may be better: it allows for more specialized training, and maximizes the interaction among students with similar comparative advantages. However, if students are uncertain about their major-specific fits, Sys.J may lead to mismatch problems. Efficiency comparisons across these two admissions systems depend on the degree of uncertainty faced by students, the importance of peer effects, and student sorting behavior that determines equilibrium peer quality. Simple cross-system comparisons are unlikely to be informative because of unobserved differences between student populations under different systems. The fundamental difficulty, that one does not observe the same population of students under two different systems, has prevented researchers from conducting efficiency comparisons and providing necessary information for policy makers contemplating admissions policy reforms. We take a first step in this direction, via a structural approach.

We develop a model of student sorting under Sys.J, allowing for uncertainties over student-major fits and endogenous peer quality that affects individual outcomes. Our first goal is to understand the equilibrium sorting behavior among students in Sys.S. Our second goal is to examine changes in student welfare and the distribution of educational outcomes if, instead of college-major-specific, a college-specific admissions regime is adopted. We apply the model to the case of Chile, where we have obtained detailed micro-level data on college enrollment and on job market returns. Although our empirical analysis focuses on the case of Chile, our framework can be easily adapted to other countries with similar admissions systems.

[1] With the exception of Quebec province.

In the model, students differ in their (multi-dimensional) abilities and educational preferences; and they face uncertainty about their suitability for various majors. The cost of and return to college education depend not only on one's own characteristics, but may also on the quality of one's peers. In the baseline case (Sys.J), there are two decision periods. First, a student makes a college-major enrollment decision, based on her expectations about peer quality across different programs and about how well suited she is to various majors. The choices of individual students, in turn, determine the equilibrium peer quality. In the second period, a college enrollee learns about her fit to the chosen major and decides whether or not to continue her studies.

In our main set of counterfactual policy experiments (Sys.S), a planner chooses optimal college-specific, rather than college-major-specific, admissions policies; a student makes an enrollment decision, chooses her course-taking intensity across different majors in the first college period, and subsequently chooses her major. Taking into account the externality arising from peer effects, the planner's optimal admissions policy guides student sorting toward the maximization of their overall welfare.

Several factors are critical for the changes in equilibrium outcomes as Sys.J switches to Sys.S. The first factor is the degree of uncertainty students face about their major-specific fits, which we find to be nontrivial. Indeed, postponing the choice of majors increases the college retention rate from 75% in the baseline to 86% under our preferred specification of Sys.S. Even under an overly pessimistic specification, the college retention rate increases to over 78%.

Second, in contrast to Sys.J, where peer students are from the same major upon college enrollment, Sys.S features a more dispersed peer composition in first-period classes. While students differ in their comparative advantages, some students have absolute advantages in multiple majors, and some majors have superior student quality. With the switch from Sys.J to Sys.S, on the one hand, the quality of first-period peers in "elite" majors will decline; on the other hand, "non-elite" majors will benefit from having better students in their first-period classes. The overall efficiency depends on, among other factors, which of these two effects dominates. Our estimation results show that for "elite" majors, own ability is more important than peer ability in determining one's market return, while the opposite is true for "non-elite" majors, suggesting that the second effect may dominate.

Finally, as students spend time trying different majors, specialized training is delayed. Welfare comparisons vary with how costly this delay is. Average student welfare will increase by 3%, if delayed specialization under Sys.S does not reduce the amount of marketable skills one obtains in college compared to Sys.J. At the other extreme, if the first period in college contributes nothing to one's skills under Sys.S, and if all students have to make up for this loss by extending their college life accordingly, a 1% loss in mean welfare will result. In a more realistic setting, we make the extension of college life a function of a student's course-taking decision in the first period, such that a shorter extension is needed for a student who has taken more courses in her major. Under this specification, student welfare increases by 1% compared to Sys.J. Moreover, female, low-income and/or low-ability students are more likely to benefit from such a switch, at the cost the most advantaged students.

Previous literature has established non-trivial uncertainty faced by students when making schooling choices. For example, Cunha, Heckman and Navarro (2005) decomposes the variability of earnings into ex-ante heterogeneity and uncertainty. They find that uncertainty accounts for about 40% of the total variability in returns to schooling. Stange (2012) finds that 14% of the total value of the opportunity to attend college is the option value arising from sequential schooling decisions made in the presence of uncertainty and learning about academic ability.

Closely related to our paper are studies that emphasize the multi-dimensionality of human capital with the presence of uncertainty. For example, Altonji (1993) introduces a model in which college students learn their preferences and probabilities of completion in two fields of study. Arcidiacono (2004) estimates a structural model of college and major choice in the U.S. in which students learn about their abilities via test scores in college before settling into their majors. As in our paper, he allows for peer effects.[2] Focusing on individual decisions, he treats peer quality as exogenous.[3] Silos and Smith (2012) estimate a model of human capital portfolio choices by agents who know their abilities in skill acquisition but face uncertainties over their fits to different occupations. Kinsler and Pavan (2014) estimate a model with both skill uncertainty and specificity

[2] There is a large and controversial literature on peer effects. Methodological issues are discussed in Manski (1993), Moffitt (2001), Brock and Durlauf (2001), and Blume, Brock, Durlauf and Ioannides (2011). Limiting discussion to recent research on peer effects in higher education, Sacerdote (2001) and Zimmerman (2003) find peer effects between roommates on grade point averages. Betts and Morell (1999) find that high-school peer groups affect college grade point average. Arcidiacono and Nicholson (2005) find no peer effects among medical students. Dale and Krueger (2002) have mixed findings.

[3] Stinebrickner and Stinebrickner (2011) use expectation data to study student's choice of major. Altonji, Blom and Meghir (2012) provides a comprehensive survey of the literature on the demand for and return to education by field of study in the U.S.

of the return to schooling, where one determinant of wage rates is how related one's job is to his major.

While this literature has focused on individual decision problems, our goal is to study the educational outcomes for the population of students, and to provide predictions about these outcomes under counterfactual policy regimes. One cannot achieve this goal without modeling student sorting in an equilibrium framework, because peer quality may change as students re-sort themselves under different policy regimes. In its emphasis on equilibrium structure, our paper is related to Epple, Romano and Sieg (2006) and Fu (2014). Both papers study college enrollment in a decentralized market, where colleges compete for better students.[4] Given our goal of addressing efficiency-related issues, and the fact that colleges in Sys.J countries are often coordinated, we study a different type of equilibrium, where the players include students and a single planner. In this centralized environment, we abstract from the determination of tuition, which is likely to be more important in decentralized market equilibria studied by Epple, Romano and Sieg (2006) and Fu (2014). Instead, we emphasize aspects of college education that are absent in these two previous studies but are more essential to our purpose: the multi-dimensionality of abilities and uncertainties over student-major fits. Moreover, we relate college education to job market outcomes, which is absent in both previous studies.

Studies comparing across different admissions systems are relatively scarce. Ofer Malamud has a series of papers that compare the labor market consequences of the English (Sys.J) and Scottish (Sys.S) systems. Malamud (2010) finds that average earnings are not significantly different between the two countries, while Malamud (2011) finds that individuals from Scotland are less likely to switch to an unrelated occupation compared to their English counterparts, suggesting that the benefits to increased match quality are large enough to outweigh the greater loss in skills from specializing early. These findings contribute to our understanding of the relative merits of the two systems, but with the caveat that students in two countries may differ in unobservable ways. Our paper compares the relative efficiency of alternative systems for *the same population of students.*

Also related to our work, Hastings, Neilson and Zimmerman (2013) (HNZ) estimate the returns to postsecondary admissions, using regression discontinuities from the centralized admissions system in Chile. They find highly heterogenous returns by

[4] Epple, Romano and Sieg (2006) model equilibrium admissions, financial aid and enrollment. Fu (2013) models equilibrium tuition, applications, admissions and enrollment.

selectivity, field of study and course requirements. The authors conclude that frictions exist in the matching between students and postsecondary degrees, including stringent admissions cutoffs for high-return programs, non-pecuniary benefits from different programs and misinformation about program-specific returns. Our paper complements HNZ by explicitly modeling student college-major choices in an equilibrium framework, allowing for heterogenous program-specific pecuniary and non-pecuniary returns, peer effects and uncertainty over the quality of student-major matches.

The rest of the paper is organized as follows: Section 2 provides some background information about education in Chile, which guides our modeling choices. Section 3 lays out the model. Section 4 describes the data. Section 5 describes the estimation and identification. Section 6 presents the empirical results. Section 7 conducts counterfactual policy experiments. The last section concludes the paper. The appendix contains additional details and tables.

2 Background: Education in Chile

There are three types of high schools in Chile: scientific-humanist (regular), technical-professional (vocational) and artistic. Most students who intend to go to college attend the first type. In their 11th grade, students choose to follow a certain academic track based on their general interests, where a track can be humanities, sciences or arts. From then on, students receive more advanced training in subjects corresponding to their tracks.

The higher education system in Chile consists of three types of institutions: universities, professional institutes, and technical formation centers. Universities offer licentiate degree programs and award academic degrees. In 2011, total enrollment in universities accounted for over 60% of all Chilean students enrolled in the higher education system. There are two main categories of universities: the 25 traditional universities and the over 30 non-traditional private universities. Traditional universities comprise the oldest and most prestigious two universities, and institutions derived from them. They are coordinated by the Council of Chancellors of Chilean Universities (CRUCH), and receive partial funding from the state. In 2011, traditional universities accommodated about 50% of all college students pursuing a bachelor's degree.

In our analysis, colleges refer only to the traditional universities for several reasons. First, we wish to examine the consequences of a centralized reform to the admissions process. This experiment is more applicable to the traditional universities, which are coordinated and state-funded, and follow a single admissions process. Second, non-traditional private universities are usually considered inferior to the traditional universities; and most of them follow (almost) open-admissions policies. We consider it more appropriate to treat them as part of the outside option for students in our model. Finally, we have enrollment data only for traditional universities.

The traditional universities employ a single admission process: the University Selection Test (PSU), which is similar to the SAT test in the U.S. The test consists of two mandatory exams, math and language, and two additional specific exams, sciences and social sciences. Taking the PSU involves a fixed fee but the marginal cost of each exam is zero.[5] Students following different academic tracks in high school will take either one or both specific exam(s). Together with the high school GPA, various PSU test scores are the only components of an index used in the admissions process. This index is a weighted average of GPA and PSU scores, where the weights differ across college programs. College admissions are college-major specific. A student must choose a college-major pair (program) jointly. A student is eligible for a program if her test score index is above the program's cutoff. After the PSU test, a common centralized application procedure is used to allocate students to different programs. A student submits an ordered list of up to eight programs. A student is admitted to and only to the program she listed the highest among all the programs she applied for and is eligible for. Because the maximum number of programs one can apply for is smaller than the total number of programs, a student may refrain from listing a program she prefers but is ineligible for. However, it is always optimal for one to reveal her true preferences for programs she is eligible for, i.e., to list her most preferred program first, her second most preferred program second, etc. An applicant can either enroll in the program to which she is admitted or opt for the outside option. As such, we can infer that the program in which a student is enrolled is the one she prefers most among *all* programs she is eligible for and the outside option. Similarly, a student (applicant or not), who is eligible for at least one program but is enrolled in none, prefers the outside option over all programs she is eligible for. Given this logic, we model a student's choice among all programs she is eligible for and the outside option without modeling the applications procedure.[6]

[5] In 2011, the fee was 23,500 pesos (1 USD was about 485 Chilean pesos in 2011).

[6] The fact that some students were ineligible for some particular programs makes it impossible to identify their preferences for these programs non-parametrically, an issue we will discuss in the

Transfers across programs in CRUCH are rare.[7] Besides a minimum college GPA requirement that differs across programs, typical transfer policies require that a student have studied at least two semesters in her former program and that the contents of her former studies be comparable to those of the program she intends to transfer to. In reality, the practice is even more restrictive. According a report by the OECD, "students must choose an academic field at the inception of their studies. With a few exceptions, lateral mobility between academic programmes is not permitted, even within institutions. This factor, combined with limited career orientation in high school, greatly influences dropout rates in tertiary education."[8] The same report also notes that the highly inflexible curriculum design further limits the mobility between programs.[9] If a student dropped out in order to re-apply to other programs in traditional universities, she must re-take the PSU test.[10]

It is worth noting that the institutional details in Chile are similar to those in many other countries, such as many Asian countries (e.g., China and Japan) and European countries (e.g., Spain and Turkey), in terms of the specialized tracking in high school, a single admissions process and rigid transfer policies. Appendix C provides further descriptions of the systems in these other countries.

3 Model

This section presents our model of Sys.J, guided by the institutional details described above. A student makes her college-major choice, subject to college-major-specific admissions rules. After the first period in college, she learns about her fitness for her major and decides whether or not to continue her studies.

---

estimation section.

[7] The rigidity of transfer policies in private colleges differs by college quality. Top private colleges use similarly rigid transfer policies while lower-ranked private colleges use more flexible transfer policies.

[8] *Reviews of National Policies for Education: Tertiary Education in Chile* (2009) OECD, page 146.

[9] "A review of the curricular grid shows a rigid curriculum with very limited or no options (electives classes) once the student has chosen an area of specialisation. In some cases, flexibility is incorporated by making available a few optional courses within the same field of study." page 143.

[10] This was true for cohorts in our sample. A new policy was announced recently that allows students to use one-year-old PSU test results for college application.

3.1 Primitives

There is a continuum of students with different gender, family income (y), abilities (a) and academic interests. There are J colleges, each with M majors. Let (j, m) denote a program. Admissions are subject to program-specific standards. An outside option is available to all students. To save notation, we omit student subscript i.

3.1.1 Student Characteristics

A student comes from one of the family income groups $y \in \{low, high\}$,[11] has multi-dimensional knowledge in subjects such as math, language, social science and science, summarized by $s = [s_1, s_2, ..., s_S]$, the vector of test scores. Various elements of such knowledge are combined with the publicly known major-specific weights to form major-specific (pre-college) ability,

$$a_m = \sum_{l=1}^{S} \omega_{ml} s_l, \tag{1}$$

where $\omega_m = [\omega_{m1}, ..., \omega_{mS}]$ is the vector of major-m-specific weights and $\sum_{l=1}^{S} \omega_{ml} = 1$. ω_m's differ across majors: for example, an engineer uses more math and less language than does a journalist. Notice that abilities are correlated across majors as multi-dimensional knowledge is used in various majors.

Given the different academic tracks they follow in high school, some students will consider only majors that emphasize knowledge in certain subjects, while some are open to all majors. Such general interests are reflected in their abilities.[12] Let M_a be the set of majors within the general interest of a student with ability vector a.[13] Denote student characteristics that are observable to the researcher, i.e., the vector of abilities, family income and gender by the vector $x \equiv [a, y, I(female)]$, and its distribution by $F_x(\cdot)$.

[11] $y = low$ if family income is lower than the median among Chilean households.

[12] Without increasing the test fee, taking both the science and the social science exams will only enlarge a student's opportunity set. A student who does not take the science exam will not be considered by programs that require science scores, but her admissions to programs that do not require science scores will not be affected even if she scores poorly in science. However, some students only take either the science or the social science exam; we view this as indication of their general academic interests. We treat students' preferences and abilities as pre-determined.

[13] Letting $a_m = NA$ if a student does not take the subject test required by major m, M_a is given by

$$M_a = \{m \in \{1, ..., M\} : a_m \neq NA\}.$$

3.1.2 Course Bundles, Skills and Wages

The building blocks for skill formation are various categories of courses $n = 1, ..., N$. Different college-major programs use different bundles of courses, which makes the production of skills college-major-specific. Let $\rho_{jmn} \in [0,1]$ be the weight on course category n required by program (j, m). Let $\rho_{jm} = \{\rho_{jmn}\}_{n=1}^{N}$ be the vector of course weights, such that $\sum_n \rho_{jmn} = 1$ for each (j, m). The technological differences between two programs are governed by their bundlings $\left(\rho_{jm}\right)$. The vector $\{\rho_{jmn}\}_{m=1}^{M}$ reflects how general Course n is. At the two extremes, a course is completely general if $\rho_{jmn} > 0$ for all m's; and a course is purely major-specific if $\rho_{jmn} = 0$ for all but one major.

Given the course requirement $\left(\rho_{jm}\right)$ in program (j, m), a student's skill attainment depends on her own major-specific ability (a_m), the quality of her peers (A_{jm}) and how efficient she is at various courses, where peer quality A_{jm} is the average major-m ability of enrollees in (j, m).[14] A student learns how efficient she is in each course type *after* being exposed to it. Let η_n be the student's efficiency in course n. We assume that η_n's are independent of each other, each drawn from $N\left(0, \sigma^2_{\eta_n}\right)$ and that $\eta = \{\eta_n\}_n$ is i.i.d. across students.[15] Notice that although η_n's are independent, a student's fitness for different courses are correlated because her ability affects her skill formation in all courses.

The human capital production function in (j, m) is given by[16]

$$\ln h_m\left(a_m, A_{jm}, \eta, \rho_{jm}\right) = \varphi_m \ln\left(A_{jm}\right) + \sum_n \rho_{jmn}\left[\gamma_n \ln\left(a_m\right) + \eta_n\right]. \tag{2}$$

The first term summarizes the contribution of one's peers, governed by the major-specific parameter $\varphi_m \geq 0$. The second part summarizes the importance of own ability and efficiency, which is a sum of the contribution of each course required by program (j, m).

Wages are major-specific functions of one's human capital (hence of $a_m, A_{jm}, \eta, \rho_{jm}$),

[14] Peer quality may affect market returns via different channels, such as human capital production, statistical discrimination, social networks, etc. Our data do not allow us to distinguish among various channels. For ease of illustration, we describe peer quality in the framework of human capital production. Arguably, the entire distribution of peer ability may matter. For feasibility reasons, we follow the common practice in the literature and assume that only the average peer quality matters.

[15] In our empirical analyses, similar courses are categorized into one group n, which makes the assumption that η_n's are independent across course groups weaker.

[16] Notice that $h_m(\cdot)$ represents the total amount of marketable skills. As such, $h_m(\cdot)$ may be a combination of pure major-specific skill and general skill.

work experience (τ) and gender, with a transitory idiosyncratic wage shock ζ_τ. The wage for a graduate from program (j,m) is given by

$$\begin{aligned} \ln\left(w_m\left(\tau,x,A_{jm},\eta,\rho_{jm},\zeta_\tau\right)\right) &= \alpha_{0m}+\alpha_{1m}\tau-\alpha_{2m}\tau^2+\alpha_{3m}I(female) \\ &\quad +\ln(h_m\left(a_m,A_{jm},\eta,\rho_{jm}\right))+\zeta_\tau, \end{aligned} \tag{3}$$

where ζ_τ is i.i.d. normal with standard deviation σ_ζ.

Discussion We model various courses as the building blocks of human capital as a parsimonious way to capture the following important features: 1) The ex-ante uncertainties over a student's fitness for different programs are naturally correlated across programs that have overlapping course requirements. The correlation increases with the degree of course overlapping between two programs. 2) Different programs involve different degrees of uncertainty, depending on the way courses are bundled.[17] For example, a major (e.g., medicine) with a high concentration on one particular type of courses is riskier than a major (e.g., education) that bundles courses in a more diversified way. 3) As we will see in the data, course bundlings (ρ) vary mainly across majors, reflecting the fundamental differences between, for example, the training of a journalist and that of an engineer. 4) Given the same major, colleges differ in their course requirements, which, together with peer quality, leads to different returns to education across colleges.

3.1.3 Consumption Values and Costs

The per-period non-pecuniary consumption value of a program varies with gender, own ability and peer ability as follows

$$v_{jm}(x,\epsilon,A_{jm})=\overline{v}_mI(female)+\lambda_{1m}a_m+\lambda_{2m}a_m^2+\lambda_{3m}A_{jm}+\lambda_4(A_{jm}-a_m)^2+\epsilon_{jm}. \tag{4}$$

The mean major-specific consumption values for males are set to zero, and $\overline{v}_m$ is the mean major-m value for females, reflecting the possibility that some majors may appeal more to females than to males. λ_{1m} and λ_{2m} measure how consumption values in major m change with one's major-specific ability.[18] For example, an individual with higher

[17] To see this, notice that the uncertain part of the human capital production in program (j,m) is given by $\sum_n \rho_{jmn}\eta_n \sim N\left(0,\sum_n \rho_{jmn}^2\sigma_{\eta_n}^2\right)$, hence the variance depends on $\left\{\rho_{jmn}\right\}_n$.

[18] In the estimation, we restrict $\lambda_{1m}\geq 0$ and $\lambda_{2m}=\lambda_2$ for all m.

ability a_m may find it more enjoyable to study in major m and work in major-m related jobs. λ_{3m} captures the major-specific effect of peer quality A_{jm} on one's consumption value: it may be more challenging or enjoyable to have higher-ability peers, which differs across majors. The second last term allows for the possibility that it might be more/less enjoyable to have peers whose ability is similar to one's own. Finally, ϵ_{jm} represents permanent idiosyncratic tastes, drawn from the distribution $F_\epsilon(\cdot)$. Tastes are i.i.d. across all students, but each individual student's tastes are correlated across majors within a college, and across colleges given the same major.[19]

Let p_{jm} be the tuition and fee for program (j, m). The annual monetary costs of attending program (j, m) is governed by[20]

$$C_{jm}(x) = p_{jm} + \left(c_1 p_{jm} + c_2 p_{jm}^2\right) I(y = low). \tag{5}$$

We allow the same tuition level to have different cost impacts on students from low family income group $I(y = low)$ to capture possible credit constraints.

3.1.4 Timing

There are three stages in this model.
Stage 1: Students make college-major enrollment decisions.
Stage 2: A college enrollee in major m observes her efficiency shocks $\widetilde{\eta}_{jm}$ and chooses to stay or to drop out at the end of the first period in college, where

$$\widetilde{\eta}_{jm} \equiv \left\{\eta_n | \rho_{jmn} > 0\right\}_{n=1}^{N} \tag{6}$$

is the subvector of a student's efficiency levels in courses required by (j, m).[21] Student choice is restricted to be between staying and dropping out, which is consistent with the Chilean practice mentioned in Section 2. Later in a counterfactual experiment, we explore the gain from more flexible transfer policies.
Stage 3: Stayers study one more period in college and then enter the labor market. The following table summarizes the information at each decision period.

[19] Students' tastes have been shown to have major impacts on their choices of majors, e.g., Arcidiacono (2004) and Wiswall and Zafar (2014).

[20] Financial aid from CRUCH colleges to students is extremely rare.

[21] We also assume that an enrollee fully observes her efficiency in her major by the end of Stage 2 (2 years in college). Without information on student performance in college, it is infeasible to allow for gradual learning.

Information Set: Sys.J

Stage	Student	Researcher
1: Enrollment	x, ϵ	x
2: Stay/Drop out	$x, \epsilon, \widetilde{\eta}_{jm}$	x

Remark 1 *We have assumed away ex-ante unobserved ability heterogeneity due to the non-trivial complications it will create for the estimation.*[22] *Findings from previous studies suggest that biases from the omission of such heterogeneity are likely to be small. For example, in Arcidiacono (2004), results are similar with and without unobserved student types. Hastings, Neilson and Zimmerman (2013) find evidence of enrollee selection on absolute advantages but no evidence of selection on comparative advantages. Comparative advantages are the most relevant to our policy experiments, which involve mainly the re-distribution of students across different majors.*

3.2 Student Problem

This subsection solves the student's problem backwards.[23]

3.2.1 Continuation Decision

After the first college period, an enrollee in (j, m) observes her efficiency vector $\widetilde{\eta}_{jm}$, and decides whether to continue studying or to drop out. Let $V_d(x)$ be the value of dropping out, a function of student characteristics.[24] Given peer quality A_{jm}, a student's second-period problem is

[22] In particular, one would no longer observe the equilibrium peer abilities, or which equilibrium was realized among the set of potential equilibria.

[23] To ease the notation, we present the model as if each period in college lasts one year. In practice, we treat the first two years in college as the first college period in the model, and the rest of college years as the second period, which differs across majors. Students' value functions are adjusted to be consistent with the actual time framework. See the Appendix A2.1 for details.

[24] Ideally, one would model the dropout and the outside options in further detail, by differentiating various choices within the outside option, e.g., working, re-taking the PSU test and re-applying the next year, or attending an open admissions private college. Unfortunately, we observe none of these details. In order to make the most use of the data available, we model the values of the dropout and the outside options as functions of student characteristics. Given the functional form assumptions, these value functions, hence student welfare, are identified up to a constant because 1) we have normalized the non-pecuniary value of majors to zero for males and 2) a student's utility is measured in pesos and we observe wages.

$u_{jm}(x, \epsilon, \widetilde{\eta}_{jm} | A_{jm}) =$

$$\max \left\{ \left(\begin{array}{c} v_{jm}(x, \epsilon, A_{jm}) - C_{jm}(x) + \\ \sum_{\tau'=3}^{T} \beta^{\tau'-2} \left(E_\zeta w_m \left(\tau - 3, x, A_{jm}, \eta, \rho_{jm}, \zeta \right) \right) + v_{jm}(x, \epsilon, A_{jm})) \end{array} \right), V_d(x) \right\}. \tag{7}$$

If the student chooses to continue, she will stay one more period in college, obtaining the net consumption value $v_{jm}(x, \epsilon, A_{jm}) - C_{jm}(x)$, and then enjoy the monetary and consumption value of her choice after college from period 3 to retirement period $T = 45$, discounted at rate β. Let $\delta^2_{jm}(x, \epsilon, \widetilde{\eta}_{jm} | A_{jm}) = 1$ if an enrollee in program (j, m) chooses to continue in Stage 2, and 0 otherwise.

Remark 2 *We restrict $V_d(x)$ to be the same regardless of one's prior program, because we do not have the data that would allow us to identify this value at a more disaggregated level.*[25] *For a dropout, the net benefit from going to college will be captured in the in-college net consumption value minus her tuition cost. This assumption rules out the case where partial training is more useful in some majors than in others. In the current specification, these differences will be absorbed in student's major-specific preferences and effort costs.*

3.2.2 College-Major Choice

Under the Chilean system, program (j, m) is in a student's choice set if only if $a_m \geq a^*_{jm}$, the (j, m)-specific admissions cutoff. Given the vector of peer quality in every program $A \equiv \{A_{jm}\}_{jm}$, a student chooses the best among the programs she is eligible for and

[25] Aggregate statistics show that most students who drop out of CRUCH universities remain out of the higher education system. For example, among those who entered CRUCH between 2008 and 2011 and later dropped out, within two years after they dropped out, about 18% enrolled in private colleges, 12% enrolled in non-college higher education institutes (e.g., technical formation centers), and 70% were not enrolled in the higher education system at any point in those two years.

the outside option with value $V_0(x)$,[26] i.e.,

$$U(x,\epsilon|a^*,A)=\max\left\{\max_{(j,m)|a_m\geq a^*_{jm}}\left\{\begin{array}{l}\beta E_{\widetilde{\eta}_m}(u_{jm}(x,\epsilon,\widetilde{\eta}_{jm}|A_{jm}))\\ +v_{jm}(x,\epsilon,A_{jm})-C_{jm}(x)\end{array}\right\},V_0(x)\right\}. \quad (8)$$

Let $\delta^1_{jm}(x,\epsilon|a^*,A)=1$ if program (j,m) is chosen in Stage 1. For a student, the enrollment choice is generically unique.

3.3 Sorting Equilibrium

Definition 1 *Given cutoffs a^*, a sorting equilibrium consists of a set of student enrollment and continuation strategies $\left\{\delta^1_{jm}(x,\epsilon|a^*,\cdot),\delta^2_{jm}(x,\epsilon,\widetilde{\eta}_{jm}|\cdot)\right\}_{jm}$, and the vector of peer quality $A=\{A_{jm}\}_{jm}$, such that*[27]
(a) $\delta^2_{jm}(x,\epsilon,\widetilde{\eta}_{jm}|A_{jm})$ is an optimal continuation decision for every $(x,\epsilon,\widetilde{\eta}_{jm})$;
(b) $\left\{\delta^1_{jm}(x,\epsilon|a^*,A)\right\}_{jm}$ is an optimal enrollment decision for every (x,ϵ);
(c) A is consistent with individual decisions such that, for every (j,m),

$$A_{jm}=\frac{\int_x\int_\epsilon \delta^1_{jm}(x,\epsilon|a^*,A)a_m dF_\epsilon(\epsilon)\,dF_x(x)}{\int_x\int_\epsilon \delta^1_{jm}(x,\epsilon|a^*,A)dF_\epsilon(\epsilon)\,dF_x(x)}. \quad (9)$$

A sorting equilibrium can be viewed as a fixed point of an equilibrium mapping from the support of peer quality A to itself. Appendix B4 proves the existence of an equilibrium in a simplified model. Appendix A3 describes our algorithm to search for

[26] The value of the outside option and that of dropout depend on one's test scores (s) and one's family income (y), both of which are elements in x. We assume that the intercepts of outside values differ across income groups, and that the value of dropout is proportional to the value of the outside option:

$$\begin{aligned}V_0(x) &= \sum_{\tau'=1}^{T}\beta^{\tau'-1}\left[\sum_{l=1}^{L}\theta_l s^l+\theta_{01}(1+\theta_{02}I(y=low))\right],\\ V_d(x) &= \varkappa V_0(x).\end{aligned}$$

[27] A sorting equilibrium takes the admissions cutoffs as given. We choose not to model the cutoff rules under the status quo (Sys.J) because our goal is to consider a different admissions regime (Sys.S) and compare it with the status quo. For this purpose, we need to understand student sorting and uncover the underlying student-side parameters, which can be accomplished by estimating the sorting equilibrium model without modeling the cutoffs. We also need to model how the admissions policies are chosen under Sys.S, which we do in the counterfactual experiments.

equilibria, which we always find in practice.[28]

4 Data

4.1 Data Sources and Sample Selection

Our first data source is the Chilean Department of Evaluation and Educational Testing Service, which records the PSU scores and high school GPA of all test takers and the college-major enrollment information for those enrolled in traditional universities. We obtained micro-level data for the 2011 cohort, consisting of $247,360$ PSU test takers. We focus on the $159,365$ students who met the minimum requirement for admission to at least one program and who were not admitted based on special talents such as athletes.[29] Most of these students did not enroll in any of the traditional universities, i.e., they have chosen the outside option in our model. From the $159,365$ students, we draw $10,000$ students as our final sample to be used throughout our empirical analyses due to computational considerations.[30]

Our second data source is Futuro Laboral, a project by the Ministry of Education that follows a random sample of college graduates (classes of 1995, 1998, 2000 and 2001). This panel data set matches tax return information with students' college admissions information, so we observe annual earnings, months worked, high school GPA, PSU scores, college and major. For each cohort, earnings information is available from graduation until 2005. We calculated the monthly wage as annual earnings divided by the number of months worked, and the (potential) annual wage as 12 times the monthly wage, measured in thousands of deflated pesos. Then, for each major, we trimmed the calculated annual wages at the $2nd$ and the $98th$ percentiles. The two most recent cohorts have the largest numbers of observations and they have very similar observable characteristics. We combined these two cohorts to obtain our measures of abilities and wages among graduates from different college-major programs. We also use the wage information from the two earlier cohorts to measure major-specific wage

[28] Uniqueness of the equilibrium is not guaranteed. However, all equilibrium objects are observed in the data, which is a fact we use in designing our algorithm.

[29] Ineligible students can only choose the outside option and will not contribute to the estimation.

[30] Some options are chosen by students at much lower frequency than others. To improve efficiency, we conduct choice-based sampling with weights calculated from the distribution of choices in the population of $159,365$ students. The weighted sample is representative. See Manski and McFadden (1981).

growth at higher work experience levels. The final wage sample consists of $19,201$ individuals from the combined 2000-2001 cohorts, and $10,618$ from the 1995 and 1998 cohorts.

The PSU data contains information on individual ability, enrollment and peer quality, but not the market return to college education. The wage data, on the other hand, does not have information on the quality of one's peers while in college. We combine these two data sets in our empirical analysis. We standardized the test scores according to the cohort-specific mean and standard deviation to make the test scores comparable across cohorts.[31] Thus, we have created a synthetic cohort, the empirical counterpart of students in our model.[32]

The wage data from Futuro Laboral contains wage information only in one's early career, up to 10 years. To obtain information on wages at higher experience levels, we use cross-sectional data from the Chilean Characterization Socioeconomic Survey (CASEN), which is similar to the Current Population Survey in the U.S. We compare the average wages across different cohorts of college graduates to obtain measures of wage growth at different experience levels. Although they are not from panel data, such measures restrict the model from predicting unrealistic wage paths in one's later career in order to fit other aspects of the data.

We also draw information from the Indices database provided by the Ministry of Education of Chile. It contains information on college-major-specific tuition, weights $(\{\omega_{ml}\})$ used to form the admission score index,[33] the admission cutoffs $\left(\left\{a^*_{jm}\right\}\right)$, and the numbers of enrollees for multiple years. Finally, we obtain information on the program-specific course requirements $\left(\left\{\rho_{jmn}\right\}\right)$ from webpages of the CRUCH colleges. For each program, we calculate ρ_{jmn} as the total credits required for course category n divided by the total credits required by program (j,m).

[31] We standardize the test scores because of the grade inflation over years. The summary statistics of the test scores are available for multiple years, although the micro-level data are not.

[32] Given data availability, we have to make the assumption that there exists no systematic difference across cohorts conditional on comparable test scores. This assumption rules out, for example, the possibility that different cohorts may face different degrees of uncertainties over student-major match quality η.

[33] We have assumed that the weights used by the colleges are the same as the ones in Equation (1), which implies colleges or the Ministry of Education, like the students in our model, have the right beliefs about the wage equations.

4.2 Aggregation of Academic Programs

For both sample size and computational reasons, we have aggregated majors into eight categories according to the area of study, coursework, PSU requirements and average wage levels.[34] The aggregated majors are: Business, Education, Arts and Social Sciences, Sciences, Engineering, Health, Medicine and Law.[35] We also aggregated individual traditional universities into three tiers based on admissions criteria and student quality.[36] Thus, students have 25 options, including the outside option, in making their enrollment decisions.[37]

Table 1 shows some details about the aggregation of programs. The second column shows the number of colleges in each tier. The third column shows the quality of students within each tier, measured by the average of math and language scores. Treating each college-level mean score as a variable, the parentheses show the cross-college standard deviations of these means within each tier. The fourth columns shows that the mean and the standard deviation of college size (total enrollment) within each tier. The average college size decreases as one goes from Tier 1 to Tier 3. However, in terms of total capacity at the tier level, Tier 2 is the largest and Tier 1 is the smallest. The last column shows the distribution of tuition levels. Cross-tier differences are clear: higher-ranked colleges have better students, larger enrollment and higher tuition. Throughout our empirical analyses, a program refers to the aggregated (tier, major).

We divide all courses into 13 categories. Eight are purely major-specific, each consisting of advanced courses required only for students in the relevant major. The other five categories are each required by at least two majors. Table A1.4 presents details about major-specific course requirements.

[34] Although we can enlarge the sample size of the PSU data by including more students, we are restricted by the sample size of the wage data. Finer division will lead to too few observations in each program.

[35] All these majors, including law and medicine, are offered as undergraduate majors in Chile. Medicine and health are different majors: medicine produces doctors and medical researchers while health produces mainly nurses.

[36] The empirical definitions of objects such as program-specific retention rates are adjusted to be consistent with the aggregation, see Appendix A2.2 for details.

[37] As a by-product of the aggregation of programs, the assumption that students cannot transfer becomes even more reasonable because any transfer across the aggregated programs will involve very different programs.

Table 1 Aggregation of Colleges

Tier	No. Colleges	Mean Score[a]	Total Enrollment[b]	Tuition[c]
1	2	702 (4.2)	21440 (2171)	3609 (568.7)
2	10	616 (17.7)	10239 (4416)	2560 (337.2)
3	13	568 (7.2)	5276 (2043)	2219 (304.2)

[a]The average of $\frac{math+language}{2}$ across freshmen within a college.

[b]Total number of enrollees per college.

[c]The average tuition (in 1,000 pesos) across majors within a college.

[d]Cross-college std. deviations are shown in parentheses.

4.3 Summary Statistics

Table 2 shows summary statistics by enrollment status. Both test scores and graduate wages increase with the ranking of tiers. Over 71% of students in the sample were not enrolled in any of the traditional universities and only 5% were enrolled in the top tier.[38] Compared to average students, females (53% of the sample) are less likely to enroll in college and a larger fraction of female enrollees are enrolled in the lowest tier.

Table 2 Summary Statistics By Tier (All Students)

	Math[a]	Language	Log Wage[b]	Dist. for All (%)	Dist. for Female (%)
Tier 1	709 (80.9)	692 (58.5)	8.91 (0.59)	5.1	4.5
Tier 2	624 (69.0)	611 (68.9)	8.57 (0.66)	14.1	12.2
Tier 3	572 (58.8)	570 (62.4)	8.32 (0.69)	9.0	9.1
Outside	533 (67.5)	532 (67.4)	-	71.8	74.2

[a]The maximum score for each subject is 850. Std. deviations across students are in parentheses.

[b]Log of starting wage in 1000 pesos.

Table 3 shows enrollee characteristics by major. The majors are listed in the order of the observed average starting wages.[39] This ranking is also roughly consistent with that of average test scores across majors. Medical students score higher in both math and language than all other students, while education students are at the other

[38]For students not enrolled in the traditional universities, we have no information other than their test scores.

[39]See Table A1.6 for wage regressions by major. See Figure 1 for the average major-specific wages by experience levels.

extreme.[40] Comparative advantages differ across majors. For example, law and social science majors have clear comparative advantage in language, while the opposite is true for engineering and science majors. The last two columns show the fraction of students in each major among, respectively, all enrollees and female enrollees. Females are significantly more likely to major in education and health but much less so in engineering.

Table 3 Summary Statistics By Major (Enrollees)

	Math	Language	Dist. for All (%)	Dist. for Female (%)
Medicine	750 (66.0)	719 (55.5)	3.4	3.2
Law	607 (74.2)	671 (72.1)	4.6	4.8
Engineering	644 (79.7)	597 (75.4)	36.6	23.4
Business	620 (87.3)	605 (73.9)	9.9	10.5
Health	628 (58.3)	632 (64.3)	11.7	17.1
Science	631 (78.2)	606 (82.1)	8.5	8.3
Arts&Social	578 (70.7)	624 (72.4)	11.2	14.1
Education	569 (59.5)	593 (64.2)	14.0	18.6

5 Estimation

The model is estimated via simulated generalized method of moments (SGMM). For a given parameter configuration, we solve for the sorting equilibrium and compute the model-predicted moments. The parameter estimates minimize the weighted distance between model-predicted moments ($M(\Theta)$) and data moments $\left(M^d\right)$:

$$\widehat{\Theta} = \arg\min_{\Theta} \left\{ \left(M(\Theta) - M^d\right)' W \left(M(\Theta) - M^d\right) \right\},$$

where Θ is the vector of structural parameters, and W is a positive-definite weighting matrix.[41] Θ includes parameters governing the distributions of student tastes, the

[40] In majors like medicine, the quality of students remains high even in the lowest tier. As shown in Table A1.5, an average medical student in Tier 3 has higher scores than an average Tier 1 student majoring in health, science, social science or education.

[41] In particular, W is a diagonal matrix, the $(k,k)^{th}$ component of which is the inverse of the variance of the k^{th} moment, estimated from the data. To calculate the optimal weighting matrix, we would have to numerically calculate the derivatives of the GMM objective function, which may lead to inconsistency due to numerical imprecision. So we choose not to use the optimal weighting matrix. Under the current weighting matrix, our estimates will be consistent but less efficient. However, as shown in the estimation results, the precision of most of our parameter estimates is high due to the

distribution of major-specific efficiency shocks, the human capital production function, the wage function, the consumption values and costs of colleges and majors, and the values of the outside and the dropout options.

Given that the equilibrium peer quality is observed and used as target moments, we have also estimated the parameters without imposing equilibrium conditions, which boils down to an individual decision model. We deem model consistency critical for the empirical analysis we do, so we focus on the first approach because it favors parameters that guarantee equilibrium consistency over those that may sacrifice consistency for better values of the SGMM objective function.[42]

5.1 Target Moments

The combined data sets contain information on various predictions of the model, based on which we choose our target moments. The PSU data contains information that summarizes the sorting equilibrium: program-specific enrollment and peer quality (Moments $1(a)$ and $2(a)$). It also provides information critical for the identification of student preferences and costs. The wage data provides information about human capital production and wage functions. Together with college retention rates, the wage data also reflects the degree of match uncertainty faced by students. In total, we estimate 95 parameters by matching 448 moments.

1. Enrollment status:
   (a) Fractions of students across tier-major (j, m) overall, among females and among low-family-income students.
   (b) Fractions of students enrolled in (j, m) with $a_m \geq a^*_{j'm}$ where j' is a tier ranked higher than j and $a_m \geq a^*_{jm}$ guarantees that the student can choose (j', m).
   (c) Fractions of students enrolled in j with $a_m \geq a^*_{jm}$ by (j, m).
2. Ability by enrollment status:
   (a) First and second moments of major-m ability (a_m) by (j, m).

relatively large sample size.

[42] Differences between the estimates from these two estimation approaches exist but are not big enough to generate significant differences in model fits or in counterfactual experiments.

(b) Mean test scores among students who chose the outside option.

(c) Retention rates by (j, m) calculated from enrollments in the college data.

3. Graduate ability: First and second moments of major-m ability among graduates by (j, m).

4. Starting wage:

   (a) First and second moments of log starting wage by (j, m).

   (b) First moments of log starting wage by (j, m) for females.

   (c) Cross moments of log starting wage and major-specific ability by (j, m).

5. Wage growth:

   (a) Mean of the first differences of log wage by major for experience $\tau = 1, ..., 9$.

   (b) From CASEN: first difference of the mean log wage at $\tau = 10, ..., 40$.

Although the entire set of model parameters work jointly to fit the data, one can obtain some intuition about identification from considering various aspects of the data that are more informative about certain parameters than others.[43] The first major set of parameters governs student preferences for different academic programs. The enrollment choices made by students with different demographics (Moments $1\,(a)$) reveal information about the relationship between these characteristics and student preferences and costs. For example, gender-specific tastes ($\overline{v}_m$) help explain the gender-specific enrollment patterns shown in Table 3 and summarized in further details by Moments $1\,(a)$ that cannot be rationalized only by the gender ability difference we observe in the data. Similarly, in our model, students from different family income groups, who are otherwise equivalent, may choose different academic programs because 1) they may view tuition costs differently and 2) their values of the outside options may be different. To separately identify these two channels, we utilize the cross-program variation in tuition levels (24 different levels).[44] The two parameters (c_1 and c_2 in Equation (5)) that

[43]We have also conducted Monte Carlo exercises to provide some evidence of identification. In particular, we first simulated data with parameter values that we choose, treated as the "truth" and then, using moments from the simulated data, started the estimation of the model from a wide range of initial guesses of parameter values. In all cases, we were able to recover parameter values that are close to the "truth."

[44]See Altonji, Blom and Meghir (2011) for a discussion of using major-specific prices for identification.

govern Effect 1) and the one parameter (θ_{02} in $V_0(x)$) that governs Effect 2) adjust in order to rationalize the different enrollment (non-enrollment) patterns across income groups that are associated with the program-specific tuition levels (Moment 1 (a)).

Students also differ in their unobservable tastes (ϵ_{jm}). Among similar students who pursued the same major, some chose higher-ranked colleges and others lower-ranked colleges (Moments 1 (b)). This informs us of the dispersion in tastes for colleges. Similar students within the same college made different major choices (e.g., more lucrative majors vs. less lucrative ones), reflecting the dispersion of their tastes for majors (Moments 1 (c)). Together with student enrollment choices (Moments 1), the distribution of abilities within a program (Moments 2 (a)) and the ability levels among those who chose the outside option (Moments 2 (b)) are informative about the relationship between peer quality and effort costs in Equation (4). For example, if high peer quality increases or barely decreases one's non-pecuniary utility (λ_{3m}), then more students who are eligible, including those who are marginally eligible, will be drawn to programs with better peers in order to benefit from the positive peer effects on wages, which will increase the ability dispersion within each of these programs. Similarly, too strong a relationship between peer quality and effort cost (λ_4) will decrease the ability dispersion within a program.

The second major set of parameters governs match uncertainty (σ_η), human capital production (2) and the wage function (3). Two assumptions greatly facilitate our identification: 1) student's pre-college abilities are observable and 2) student tastes (permanent and unobservable) are uncorrelated with student post-enrollment shocks, and they do not affect market returns. Given these two assumptions, the systematic differences in wages among similar workers (Moments 4) arise from their post-enrollment efficiency shocks: all else equal, a higher dispersion in efficiency shocks would lead to a higher dispersion of wages. College retention rates (Moments 2 (c)) are a second major source of information for identifying efficiency shocks. A lower dispersion in efficiency shocks would lead to higher retention rates. Student ability distributions conditional on choices are also informative. In particular, the likelihood that shocks are bad enough for students with relatively high pre-college ability to drop out will be low if the dispersion of efficiency shocks is low. In contrast, a highly dispersed distribution of shocks will lead a non-trivial fraction of high-ability students to drop out. These effects are directly reflected in the ability distribution among graduates relative to that among all enrollees (Moments 2 (a) and 3).

The relationship between wages and student's observable characteristics (Moments $4(b)$ and $4(c)$) provides key information about the importance of these characteristics in the human capital production and wage functions. In particular, the importance of pre-college ability is mainly captured by the correlation between wages and ability levels. Correlations are not directly targeted but they are jointly captured by the Moments $4(a)$ and $4(c)$.[45]

Finally, Moments 5 inform us of wage growth over the life cycle. Moments $5(a)$ contain major-specific early-career wage information from Futuro Laboral. Moments $5(b)$ contain information at higher experience levels. The weakness of the CASEN data is that we do not observe college major. As such, the Futuro Laboral data is the main source for us to identify different lifetime wage paths across majors, while the CASEN data helps restrict the wage path in later years over all college majors.

Remark 3 *Like many other constrained choice models, our model is not non-parametrically identified. For example, for students who are ineligible for some programs, it is not possible to non-paramEterically identify their preferences for those programs. For this reason, we have to impose assumptions that allow us to "extrapolate," including 1) student tastes for programs* ϵ_{jm} *and efficiency shocks* η *are i.i.d. across students, independent of observables* x*, and independent of each other; 2) the consumption value of a program and the human capital production function are both continuous functions of student characteristics.*

6 Results

6.1 Parameter Estimates

This section reports the estimates of parameters of major interest. Tables A2.1-A2.6 in the appendix report the estimates of other parameters. Standard errors (in parentheses) are calculated via bootstrapping.[46] Table 4 displays the roles of peer quality and own ability in the human capital production for each major, which also measure the elasticities of wages with respect to peer ability and own ability. The left panel shows

[45] Notice that Moments 4 are at the program level, the cross-program differences arise from their student quality and course requirements, both observed in the data.

[46] Calculating standard errors via standard first-order Taylor expansions might be problematic because we have to use numerical method to calculate the derivatives of our GMM objective function. We took 500 bootstrap iterations. Given the sample size (10,000) and the sampling scheme described in Footnote 38, the precison of most of our estimates is high.

the parameter estimates and the standard errors of the coefficients for peer quality, i.e., φ_m in (2). As shown in (2), the total contribution of own ability is a weighted sum of its course-specific contributions, governed by $\sum_n \rho_{jmn}\gamma_n$, where $\{\rho_{jmn}\}$ is the tier-major-specific course requirement weights from the data and $\{\gamma_n\}$ is the vector of production technology parameters. Instead of showing the estimates of $\{\gamma_n\}$, which are shown in Table A2.1, the right panel of Table 4 shows the overall importance of own ability for a given major that is comparable with that of peer quality, i.e., the cross-tier averages of $\sum_n \rho_{jmn}\gamma_n$.[47]

Table 4 Human Capital Production

	Peer Ability φ		Own Ability $\sum_n \rho_{mn}\gamma_n$
	Estimate	Std. Err.	Cross-Tier Average
Medicine	0.01	(0.01)	0.08
Law	0.25	(0.02)	1.40
Engineering	0.66	(0.02)	1.40
Business	2.04	(0.01)	1.05
Health	0.71	(0.03)	0.07
Science	2.32	(0.02)	0.31
Arts&Social	1.09	(0.01)	1.02
Education	1.23	(0.02)	1.16
Efficiency Shock σ_η	0.89	(0.01)	

The left panel of Table 4 shows significant differences in the importance of peer ability across majors: the elasticity of wage with respect to peer quality is high in business and science, and close to zero in medicine.[48] This finding is consistent with those found in previous studies. For example, HNZ find that among the eight fields they consider, the return to major is significantly higher if one were admitted to a more selective (higher-peer-quality) college than a less selective college for social science, science, business and health, but not for other fields.

Considering both the left and the right panels, we find that the relative importance of peer ability versus own ability differs systematically across majors although

[47] Differences exist in $\sum_n \rho_{jmn}\gamma_n$ across tiers, but they are small because course requirements $\{\rho_{jmn}\}_n$ are similar across tiers for a given major.

[48] As mentioned earlier, our model is silent about why peer ability affects market returns. The reasons are likely to differ across majors. For example, the high elasticity of wages with respect to peer quality in business may arise because the social network one forms in college is highly valued in the business profession.

no restriction has been imposed in this respect. In the three majors with the highest average wages, the elasticity of wage with respect to peer ability is at most half of that with respect to own ability.[49] For the relatively lower-paying majors, peer ability is more important than own ability in determining wages.[50] Similar results have been found in previous studies. For example, results from Arcidiacono (2004) indicate that the importance of own SAT scores dominates that of the peers' SAT scores in high-paying majors, while the opposite is true in low-paying majors. The major-specific relative importance of peer quality versus own ability has important implications for welfare analysis as Sys.J switches to Sys.S, because the quality of first-period peers will decline for "elite" majors, while increase for "non-elite" majors. Table 4 suggests that the former negative effect is likely to be small, while the latter positive effect may be significant.[51]

The last row of Table 4 shows the dispersion of course-specific efficiency shocks.[52] To understand the overall impacts of such uncertainty, imagine reducing the dispersion by 25%, which would increase the overall college retention rate from 75% in the baseline to 83%. Clearly, students face non-trivial uncertainty.[53] Moreover, as mentioned in the model section, uncertainty differs across programs depending on how diversified the course requirements are. Majors like medicine and law involve higher degrees of uncertainty than other majors because course bundlings in professional majors are more concentrated (Table A1.4). However, this does not mean that one should expect to see higher dropout rates in professional majors, because students in these majors have higher academic abilities and have more to gain from college.

[49] It may be surprising to see small effects of both own ability and peer ability in medicine. One possible reason is that compared to their counterpart from lower-tier medical schools who have lower pre-college ability, a higher fraction of graduates from top medical schools work in research/education-related jobs and/or in the public sector, where wages are lower than those in the private sector.

[50] One possible explanation for this pattern is labor market statistical discrimination. For example, in law and medicine, the practice of licensing and residency/internship reduces the need for statistical discrimination, making peer quality less important than one's own ability. Yet, for majors like education and general science, where individual quality is hard to determine, employers may need to rely more on statistical discrimination.

[51] Results in Table 4 are qualitatively consistent with those from data wage regressions (Table A1.6).

[52] We cannot reject the hypothesis that σ_{η_n}'s are the same across courses, and therefore choose the more parsimonious specification with a common σ_η.

[53] There are cases where our estimate of σ_η may overstate the degree of uncertainty over student-major matches. The first is the existence of some unobserved component of student ability not captured by their test scores that leads to permanent wage dispersions across workers. A second case is when there exists other post-enrollment shocks that cause a student to drop out besides the efficiency shocks.

Table 5 reports parameter estimates for major-specific consumption values. Consumption values increase most with own ability in the three majors with the highest average wages, followed by social science. However, there is no significant relationship between consumption values and own ability in other majors. The second column shows the relationship between consumption values and peer quality. The effect of high-ability peers on consumption value is negative in all majors except for engineering and science majors. The negative impact is greatest in law, followed by medicine and then education. Our model is silent on why peers have different impacts on one's consumption values across majors. Yet, the results are not unreasonable. For example, it may be costly to have high-quality peers in law programs, because students are constantly placed in competitive situations. In contrast, engineering students often need to collaborate for joint projects, which may be more enjoyable with high-quality peers. Empirically, these differential utility costs help explain why some eligible students chose other majors despite of the expected high wage in majors like law and medicine. Similarly, high peer costs in majors like education and social science help explain why many students who were above the higher-tier cutoff chose the lower-tier program. This is especially true in the choice between the second and the third tiers for those majors, where the two tiers have similar cutoffs (Table A1.2).

Table 5 Consumption Values (Major-Specific Parameters)

	Own Ability		Peer Ability		Female	
Medicine	12.99	(0.28)	-8.57	(0.18)	-2248.4	(167.9)
Law	6.26	(0.14)	-11.95	(0.30)	-2644.6	(225.5)
Engineering	4.82	(0.11)	3.76	(0.14)	-2669.7	(115.1)
Business	0.02	(0.03)	-2.44	(0.08)	347.1	(85.8)
Health	0.05	(0.06)	-2.44	(0.05)	4578.0	(72.5)
Science	0.02	(0.01)	0.33	(0.10)	-529.6	(109.1)
Arts&Social	1.77	(0.02)	-3.53	(0.10)	-963.3	(194.6)
Education	0.02	(0.04)	-5.16	(0.12)	4372.8	(60.1)

The last column of Table 5 shows that compared to males, an average female has significantly higher tastes for the conventionally "feminine" majors, i.e., health and education, slightly higher preference for business, but lower tastes for all other majors. Empirically, these taste parameters help to explain the different enrollment patterns across genders as in Table 3.[54] In Appendix B1, we show that when females are endowed with the same preferences as males, there will no longer exist majors that are obviously dominated by one gender. However, the difference in comparative advantages across genders also plays a nontrivial role in explaining their different enrollment patterns.[55]

6.2 Model Fit

Overall, the model fits the data well. Table 6 shows the fits of enrollment by tier, for all students and for females.[56] The model slightly underpredicts the fraction of students enrolled in the top tier. Table 7 shows the distribution of enrollees across majors. The fit for the distribution among all enrollees is very close, except an underprediction in the fraction of law students. For female enrollees, the model underpredicts the fractions in social sciences and law. Table 8 shows the fits of average student ability and retention rates by tier. Table 9 shows the same fits but by major.[57] The fits are good overall, but the retention rate is over-predicted for Tier 3.

Table 6 Enrollment by Tier (%)

	All		Females	
	Data	Model	Data	Model
Tier 1	5.1	4.6	4.5	3.9
Tier 2	14.1	14.0	12.2	12.9
Tier 3	9.0	9.4	9.1	9.8

[54] The model explains this pattern via three channels: 1) the distributions of pre-college ability are different across genders as seen in the data; 2) gender is allowed to enter the wage function directly; 3) preferences may differ across genders.

[55] The importance of gender-specific preferences has been noted in the literature. For example, Zafar (2009) finds that preferences play a strong role in the gender gap of major choices in the U.S.

[56] The fits for students with low family income are in the appendix Tables B3-4.

[57] The retention rates reported seem to be high for two reasons. First, we focus on the traditional colleges, which are of higher quality than private colleges. Second, consistent with our data aggregation, a student is said to have stayed in (j, m) if she stayed in any specific program within our (j, m) category.

Table 7 Enrollee Distribution Across Majors (%)

	All		Females	
	Data	Model	Data	Model
Medicine	3.4	3.6	3.2	3.0
Law	4.6	3.5	4.8	3.1
Engineering	36.6	36.5	23.4	25.1
Business	9.9	10.4	10.5	11.0
Health	11.7	11.3	17.1	17.0
Science	8.5	9.1	8.3	8.9
Arts&Social	11.2	10.8	14.1	11.9
Education	14.0	14.8	18.6	19.4

Table 8 Ability & Retention (by Tier)

	Ability[a]		Retention (%)	
Tier	Data	Model	Data	Model
1	701	697	79.3	80.3
2	624	624	76.5	74.5
3	581	578	68.1	72.5

[a]The average of major-specific ability across majors in each tier.

Table 9 Ability & Retention (by Major)

	Ability[a]		Retention (%)	
	Data	Model	Data	Model
Medicine	738	730	87.6	88.0
Law	658	632	81.3	79.0
Engineering	623	623	71.8	74.7
Business	619	618	74.6	73.0
Health	641	638	79.8	79.5
Science	622	613	63.7	63.8
Arts&Social	612	601	74.3	74.4
Education	590	590	77.1	76.4

[a]Average major-specific ability a_m in each major m.

Excluded from our target moments, the next set of fit statistics utilizes the stringent cutoff rules in Chile to examine students' preferences for majors. In particular, we contrast the fraction of students who opted for the same major m in a lower tier when they missed the cutoff for a higher-tier program (j, m) by no more than 10 points. Table 10 shows the results. For example, the first row shows that in the data, 12% of the students who just missed the cutoff for medicine in a higher tier opted for the same major in a lower tier, while this fraction is 13.5% in the model. Although these are a small group of students, the model fit is good in general. A second set of fits, presented in Appendix Table B5, is the correlation between starting wage levels and the major-specific ability of graduates from each (tier, major). In most programs, the correlation is small and positive; in some cases, it is small and negative. Overall, the model is able to capture the patterns well.

Table 10 Pursue the Same Major (Marginally Ineligible for a Higher Tier)

(%)	Data	Model
Medicine	12.0	13.5
Law	5.8	8.4
Engineering	24.8	18.6
Business	4.6	7.9
Health	14.0	11.6
Science	5.4	5.9
Arts&Social	8.0	3.9
Education	8.4	10.1

7 Counterfactual Policy Experiments

We first introduce the counterfactual admissions regime Sys.S, providing overall cross-system comparisons. Next, we conduct a milder policy change that allows students one chance to switch programs within Sys.J. We then examine the effects of admissions systems in detail, focusing on the contrast between the baseline Sys.J and Sys.S. Finally, we check how robust our results are when human capital rental rates vary with the supply of college graduates. We focus on short-run equilibrium and take the distribution of student preferences and pre-college ability as fixed.

7.1 Overall Comparison

7.1.1 Sys.S

Under Sys.S, students choose their majors after they learn about their fits. We solve a planner's problem of maximizing total student welfare by setting college-specific, rather than college-major-specific, admission policies.[58] The constraints for the planner include: 1) a student eligible for a higher-tier college is also eligible for colleges ranked lower, and 2) the planner can use only ability a to distinguish students. These two restrictions keep our counterfactual experiments close to the current practice in Chile in dimensions other than the college-specific versus college-major-specific admissions. Restriction 1 prevents the planner from assigning a student to the college that the planner deems optimal, which is both far from the current Chilean practice and also may lead to mismatches due to the heterogeneity in student tastes. Restriction 2 rules out discrimination based on gender or family income.

There are four stages in this new environment:

Stage 1: The planner announces college-specific admissions policies.

Stage 2: Students make enrollment decisions, choosing one of the colleges they are eligible for or the outside option. An enrollee chooses course taking intensity across majors. Let $M(x, \epsilon, j, A_j)$ denote the set of majors to take course in chosen by a College-j enrollee (x, ϵ), which we specify below.

Stage 3: An enrollee learns her efficiency levels in courses she has taken, which are $\cup_{m \in M(x,\epsilon,j,A_j)} \widetilde{\eta}_{jm}$, i.e., the union of the course efficiency levels in each major within $M(x, \epsilon, j, A_j)$. Given the additional information, she chooses her final major from this set or drops out.

Stage 4: Stayers spend one more period studying in the final major of choice and then enter the labor market.

Information and Decision: Sys.S

Stage 1		Stage 2		Stage 3	
Info	Planner	Info	Student	Info	Student
a	Admissions	x, ϵ	College(j),$M(x, \epsilon, j, A_j)$	$x, \epsilon, \cup_{m \in M(\cdot)} \widetilde{\eta}_{jm}$	$\{m \in M(\cdot)$,Dropout$\}$

[58] The planner takes into account tuition and effort costs for students. To maximize social welfare, one would also include other costs of college education, for example, costs for colleges that are not fully covered by tuition revenue. This will be a relatively straightforward extension yet one that requires information that is unavailable to us.

The planner acts as the Stackelberg leader in this game. Instead of simple unidimensional cutoffs, optimal admissions policies will be based on the whole vector of student ability a. To calculate the benefit of admitting a student of ability a to a certain set of colleges, the planner has to first form an expectation of the student's enrollment, course-taking and major choices, integrating out student characteristics and tastes that are unobservable to the planner, and the efficiency shocks. Then, the planner calculates the expected value for this individual and her effect on her peers. Peer quality matters because it affects both the market return and the non-pecuniary consumption value. Overall, the planner's optimal admissions policies lead student sorting toward the maximization of total student welfare. Appendix B3 contains formal theoretical details.

Further Specifications Under a typical Sys.S, students may take courses that are more related to some majors than to others in the first period based on their preferences and expectations. To capture this fact and yet still keep the exercise feasible, we assume that in the first period, an enrollee will consider the probability that she may finally choose each of the majors if she has been exposed to all of them. Then, she will divide her time and effort across majors/departments according to these probabilities. In particular, if a student i expects that, conditional on staying in college in the second period, she will choose m with probability π_{im} given her (x, ϵ) and peer quality, where the expectation is taken over $\{\widetilde{\eta}_{jm'}\}_{m'=1}^{M}$, then her course-taking intensity (the fraction of time spent) in major m during the first period is equal to π_{im}.

To compare welfare, one factor that deserves special attention is the potential loss of major-specific human capital due to the delay in specialized training.[59] The data we have does not allow us to predict the exact change in human capital associated with the shift of admissions regimes because we do not observe the return to partial college education or student performance in college. However, it is still informative to provide bounds on welfare gains under Sys.S by considering various possible scenarios. In this paper, we explore two different sets of scenarios. In the first, we assume that to make up for the first period (2 years) of college spent without specialization, all students have to spend, respectively, 0, 1 and 2 extra year(s) in college. The two extreme cases may provide bounds for the welfare under Sys.S. In a second (more reasonable) scenario, the extra time one needs to spend in college is a function of her time allocation in

[59] On the other hand, if the labor market values the width of one's skill sets, one would expect greater gains from the new system than those predicted in this paper.

the first period, given by $ext(\pi)$. To reflect the fact that colleges are normally not fully flexible and only provide two graduation seasons per year, we assume that the extra time in college is measured in multiples of semesters (0.5 years), and that it is a step function.[60] Notice that the extension period is a function of student choice $\{\pi_{im}\}_m$, which in turn depends on the extension time. An equilibrium requires that student choices be self-consistent.[61]

Remark 4 *One of the choices available under the second scenario of Sys.S is to fully specialize from the beginning, which is the only choice available for a college student under Sys.J. However, it does not imply that a student will be better off, because equilibrium peer quality will change when the system changes.*

Results Table 11 shows the equilibrium enrollment, retention and student welfare under the baseline and under Sys.S for the two sets of scenarios. In all cases, postponing major choices increases the overall retention rate: a significant fraction of dropouts occur in the current system because of student-major mismatches.[62] In the no-extra-time case, enrollment increases from 28% to 36%, retention rate increases from 75% to 88%, and the mean student welfare increases by about 3.7 million pesos or 3%. When one has to spend extra time in college, enrollment decreases sharply. In the extreme case where one has to spend 2 more years in college, the new system causes a 1% welfare loss relative to the baseline. In the endogenous extension case, welfare improves by 1%.

Table 11 Enrollment, Retention & Welfare: Sys.S

	Baseline	Extra Years in College			
		0 Year	1 Year	2 Years	$ext(\pi)$
Enrollment (%)	28.0	36.2	30.7	27.2	31.4
Retention (%)	74.9	87.9	82.9	78.3	85.6
Mean Welfare (1,000 pesos)	146,495	150,166	146,997	145,162	147,834

[60] See Appendix B2.4 for details.

[61] In terms of algorithm, for every simulated student, we have an inner loop that searches for a fixed point $\pi_i = \{\pi_{im}\}_m$ that maintains this self-consistency, i.e., the solution to

$$\pi_{im} = \Pr(m \text{ is chosen by } i \text{ among the } M \text{ majors in the 2nd period} | ext(\pi_i)), \text{ for all } m.$$

[62] If students face uncertainties besides efficiency shocks, we might over-predict the retention rate in the new system.

7.1.2 Rematch Under Sys.J

Although the same rigid transfer policies are practiced in many countries like Chile, some other countries (e.g., England) with the same admissions system are more flexible in terms of transfers. To explore how much can be gained from such flexibility, the following policy experiment allows students under Sys.J one chance to rematch after the first period in college. The timing under this policy is:

Stage 1: Students make college-major enrollment decisions, subject to college-major-specific admissions policies.[63]

Stage 2: A college enrollee in major m observes her efficiency vector $\widetilde{\eta}_{jm}$, and chooses to stay, to transfer to a different program, or to drop out at the end the first period in college. To prevent arbitrage, we impose the same admissions standards on transfers.

Stage 3: Students who chose to stay in Stage 2 stay one more period in college and then enter the labor market. Transfer students observe their fits for their new majors and decide whether to stay and later enter the labor market or to drop out.

We consider three cases where a transfer student has to spend 0, 1 or 2 extra year(s) in college, compared to a non-transfer student. Under the rematch policy, the enrollment rate is over 30% even if transfer students have to spend 2 more years in college. However, the final retention rate remains similar to that in the baseline under the no-extra-time scenario and lower than the retention rate in the baseline when extra time is involved. The opportunity to rematch encourages more students to enroll but is not very effective keeping them in college. This is true despite the fact that in Stage 2 when a student makes transfer decisions, she already has partial knowledge about her fits to other majors for which the course requirements overlap with her first major.

Table 12 Enrollment, Retention & Welfare: Sys.S vs. Rematch

	Baseline	0 Extra Year	1 Extra Year	2 Extra Years
Enrollment (%)	28.0	35.8	33.1	30.0
Retention (%)	74.9	75.0	73.7	73.2
Mean Welfare (1,000 Peso)	146,495	149,064	147,938	147,067

7.2 A Closer Look

We will focus on the endogenous extra time specification ($ext(\pi)$) in our following analyses of Sys.S.

[63] We impose the same admissions policies used in the current Chilean system and re-solve the sorting equilibrium. Results from this experiment are subject to these exogenous admissions policies.

7.2.1 Gainers and Losers

The impacts of a change from Sys.J switches to Sys.S differ across students. Table 13 presents the outcomes by quartiles of test scores (math+language). Enrollment rates increase in all three lower quartiles, especially for the lowest quartile; the highest quartile, in contrast, sees a decline in enrollment rates.[64] Retention rates improve for all groups.[65] Average welfare improves for students in the first three quartiles, especially the middle two quartiles, while it decreases for the highest quartile. To further investigate who is likely to gain/lose, we run a regression of one's winner/loser status against one's characteristics (Appendix B3). We find that females and students from low income families are more likely to be gainers, and that when a student already has a clear comparative advantage as reflected in pre-college abilities, the cost of delayed specialization is likely to outweigh its benefit. All these findings suggest that Sys.J favors more advantaged students and that a switch into Sys.S would improve equity.

Table 13 Outcome by Test Score Quartiles

	<=1st Qua.	1st~2nd Qua.	2nd~3rd Qua.	>3rd Qua.
Enrollment (%)				
Baseline	1.8	11.0	37.1	62.1
$ext(\pi)$	7.6	20.3	43.0	54.9
Retention (%)				
Baseline	83.9	78.3	72.1	76.0
$ext(\pi)$	87.3	85.8	85.2	85.6
Mean Welfare (1,000 pesos)				
Baseline	129,034	133,396	150,629	173,118
$ext(\pi)$	129,856	136,993	152,666	172,004

Test score: (math+language)

7.2.2 Enrollment and Major Choice Distribution

Table 14 displays enrollment and retention rates by tier. Compared to the baseline case, Sys.S features more students enrolled in every tier and a higher fraction enrolled in the

[64] Although not targeted, the enrollment rates under the baseline match the data well, where the four rates are predicted as $[2.6, 15.3, 32.6, 62.4]$. We do not have data on retention rates by ability groups.

[65] Notice that the retention rate for the lowest quartile is the highest. This is because those who have low scores and choose to enroll are a highly selected group of students who have strong preferences for college.

top tier. Under the baseline, a nontrivial fraction of students were eligible to enroll in Tier 1 but only for majors other than their ex-ante most desirable ones. Among these students, some opted for their favorite majors in lower tiers rather than a different major in Tier 1. Under Sys.S, the planner still deems (some of) these students suitable for Tier 1, and some of them will matriculate. This is because, regardless whether or not these students eventually choose their ex-ante favorite majors, given their relatively high ability, enrolling them in Tier 1 does not have a significant negative effect on peer quality, while the improved match quality significantly increases the benefit of doing so. Retention rates in all three tiers improve significantly with the change of the system.

Table 14 Enrollment and Retention (%)

	Baseline		$ext(\pi)$	
	Enrollment	Retention	Enrollment	Retention
Tier 1	4.6	80.3	5.9	88.1
Tier 2	14.0	74.5	15.7	85.0
Tier 3	9.4	72.5	9.8	84.9

Table 15 displays the distribution of students across majors in the first and second period in college. Focusing on the first four columns, we see that without major-specific barriers to enrollment, the number of students majoring in law and medicine increases significantly compared to the baseline level. Business and health also grow.

Table 15 Distribution Across Majors (%)

	Baseline		$ext(\pi)$		Rationed $ext(\pi)$	
	1st Period	2nd Period	1st Period	2nd Period	1st Period	2nd Period
Medicine	1.0	0.9	-	1.6	-	1.0
Law	1.0	0.8	-	1.6	-	1.0
Engineering	10.2	7.6	-	8.4	-	8.5
Business	2.9	2.1	-	3.1	-	2.9
Health	3.2	2.5	-	4.3	-	3.2
Science	2.6	1.6	-	1.5	-	1.7
Arts&Social	3.0	2.3	-	2.3	-	2.4
Education	4.1	3.2	-	4.0	-	4.0
All	28.0	21.0	31.4	26.9	31.0	24.7

For the 1st period, the distribution across majors is defined only for the baseline, since in the new system students do not declare majors until the 2nd period.

Capacity Constraints and Rationing In the short run, it may not be feasible to expand each academic program to accommodate the demand. As such, we impose program-specific cutoffs in the second period such that only students above the cutoffs can choose a particular program. The cutoffs are imposed on a student's quality taking into account her realized efficiency levels.[66] We conduct a series of experiments to search for the lowest necessary set of cutoffs, such that the final number of students in each program does not exceed the enrollment size of the corresponding program under the baseline. The result under this final set of cutoffs is shown in the last two columns. Compared to the case without cutoffs, there is a small inflow to engineering, science and social science from other majors with binding cutoffs; yet all these three majors remain under capacity as many students choose either to drop out or not to enroll instead of opting for a different major. Compared to the baseline, however, every major graduates more students.

7.2.3 Productivity

Next we compare major-specific productivity across the two systems. This comparison is unclear ex ante because conflicting factors coexist, some of which include 1) Sys.S provides students with information on their match quality before they choose their majors, which should improve individual productivity ceteris paribus, 2) a switch to Sys.S improves peer quality for some majors and decreases it for other majors, and 3) under Sys.S, the pool of college graduates expands and some lower-ability students flow into majors that they were not eligible under the baseline. The results are in Table 16, which shows the average productivity as measured by mean log starting wages (in 1,000 pesos) in each major. Without rationing, productivity improves in law, health, social science and education; while it decreases in the other four majors. Rationing improves the average productivity in all majors relative to the case without rationing. In particular, with binding cutoffs, rationing reverses the ranking between Sys.J and Sys.S in terms of the productivity in medicine. However, the average productivity in business, engineering and science remain lower under Sys.S than their baseline levels.[67]

[66] In particular, a student i is evaluated based on

$$\sum_n \rho_{jmn} \left[\gamma_n \ln(a_{im}) + \eta_{in} \right].$$

[67] If post-enrollment shocks do not directly enter the wage function and if they are only partly correlated with one's future productivity, then, a switch from Sys.J to Sys.S is likely to have smaller

Table 16 Log Starting Wage

	Baseline	$ext(\pi)$	Rationed $ext(\pi)$
Medicine	9.08	8.66	9.16
Law	9.18	9.48	10.0
Engineering	8.88	8.70	8.71
Business	8.60	8.42	8.45
Health	8.65	8.77	9.06
Science	8.36	8.23	8.26
Arts&Social	8.46	8.47	8.48
Education	8.06	8.08	8.23

7.3 Endogenous Human Capital Rental Rates

In our analyses so far, we have taken the rental rates for human capital as given by the parameters in the wage function $\{e^{\alpha_{0m}}\}_m$. In this subsection, we relax this assumption and check how robust our results will be when labor market returns vary with the number and the composition of college graduates.[68] We endogenize rental prices as marginal products of a nested aggregate CES production function, given by

$$\left(g_1 L^{\upsilon_1} + g_2 H^{\upsilon_1}\right)^{\frac{1}{\upsilon_1}}, \tag{10}$$

where L is the low-skilled labor measured in unit of high-school equivalent,[69] H is the aggregate human capital among college-educated workers, and $\sigma_{e_1} = \frac{1}{1-\upsilon_1}$ is the elasticity of substitution between low-skill and high-skill human capital. For the aggregation of skilled labor across different majors, we assume a CES functional form

$$H = \left(\sum_{m=1}^{M} g_{3m} h_m^{\upsilon_2}\right)^{\frac{1}{\upsilon_2}}, \tag{11}$$

where $\{h_m\}_{m=1}^{M}$ is the vector of aggregate major-specific human capital, $g_{3m} > 0$, $\sum g_{3m} = 1$. $\sigma_{e_2} = \frac{1}{1-\upsilon_2}$ is the elasticity of substitution between different major-specific

effects on productivity.

[68]This exercise is a robustness check. A more comprehensive model that accounts for broader general equilibrium features will be a direction for future research.

[69]Following Card (2009), we assume that each worker without a high school degree supplies 0.7 units of high-school labor, and that each worker with some college education supplies 1.2 units of high-school labor.

skills.

Our data do allow us to back out all the parameters in (10) and (11) . Instead, we will consider a range of $\nu_1 \in [0.3, 0.7]$, i.e., $\sigma_{e_1} \in [1.43, 3.33]$ that has been estimated in the literature, and a wide range of υ_2 including Cobb-Douglas and linear.[70] Given each pair of (υ_1, υ_2), baseline skill levels $\{L^0, \{h_m^0\}_m\}$ and their baseline rental rates,[71] we can solve for the rest of the parameters $\{g_1, g_2, \{g_{3m}\}_m\}$ by setting the rental rates equal to the marginal products, which gives us one set of (g, υ) .

Given each set of (g, υ), we embed the endogenous major-specific human capital rental rates into our equilibrium Sys.S model. A change in the admissions system may lead to several changes to the composition of labor supply: 1) an increase in college enrollment will reduce the supply of high-school-educated workers; 2) a change in the number of college dropouts will change the number of workers with some college education; and 3) changes in the number and the distribution of college graduates across majors will change the composition of the high-skilled labor force.[72] The rental rates of various types of skills will vary accordingly, which will in turn affect the decisions made by the social planner and the students. An equilibrium requires that the rental rates are consistent with these decisions.

To focus on the short run, which is what our model is tailored for, we consider changes in the flow of one or two recent cohorts of workers into the labor market, while holding the stock of older cohorts fixed. Again, we focus on Sys.S with the endogenous extension time $ext(\pi)$ without rationing. Comparing the results from the previous subsection with their counterparts when rental rates are endogenized under different sets of aggregate production parameters, we find that our results are robust. To save space, we will report the case with the most significant changes in human capital rental rates, especially for the two fastest-growing majors, law and medicine. This happens when two (instead of one) cohorts are put under Sys.S and when $(\nu_1, \nu_2) = (0.7, 0)$, i.e., when low skill and high skill are highly substitutable while different major-specific

[70] For example, σ_{e_1} is reported to be 1.786 in Acemoglu (2002), between 1.43 and 3.33 in Ottaviano and Peri (2008), around 1.54 in Goldin and Katz (2008), between 1.5 and 2.5 in Card (2009). For υ_2, to the best of our knowledge, there exists no estimate of the elasticity of substitution across different college majors.

[71] The baseline rental rates are the estimated $\{e^{\alpha_{0m}}\}_m$ .The baseline L^0 and its rental rate are calculated using CASEN data on labor supply and wages by education group. The baseline $\{h_m^0\}_m$ is predicted by the baseline model and aggregated across experience groups. Details available upon request.

[72] In calculating these changes, we assume that all college students are represented in our model, not just those in the CRUCH colleges. This is likely to overstate the changes in labor supply and rental prices. It will be reassuring if our results are robust even to these overstated changes.

high skills aggregate in a Cobb-Douglas fashion.[73] Table 17 contrasts the outcomes under this specification with those under the partial equilibrium case with fixed rental rates. In particular, it shows the major-specific percentage changes, relative to the partial equilibrium case, in the average starting wage and the measure of graduates for the recent cohort. Medicine sees the biggest percentage decrease in the average starting wage by 3.9% and the number of graduates by 4.6%. Law experiences similar decreases but to a lesser extent. All other majors experience a slight increase in both outcomes. Endogenous skill prices do affect outcomes in the short run, but only very slightly. As a result, the average student welfare is almost the same as in the partial equilibrium case.

Table 17 Changes from Partial to GE (%)

	Mean Starting Wage	Graduates
Medicine	-3.9	-4.6
Law	-0.8	-2.3
Engineering	0.8	0.5
Business	1.6	1.8
Health	2.2	2.4
Science	1.0	0.2
Arts&Social	0.4	0.4
Education	2.8	2.1

8 Conclusion

College-major-specific admissions system (Sys.J) and college-specific admissions system (Sys.S) both have their advantages and disadvantages: whether or not the total welfare of students under one system will improve under the alternative system becomes an empirical question, one that has significant policy implications. However, answering this question is very difficult since one does not observe the same population of students under both regimes. In this paper, we have taken a first step.

We have developed and estimated an equilibrium college-major choice model under Sys.J, allowing for uncertainty and peer effects. Our model has been shown to match

[73] Intuitively, when different major-specific skills are less substitutable, an imbalance of major composition will have a larger negative impact on the level of aggregate high skill, ceteris paribus. This effect will be stronger when low skill is a better substitute for high skill.

the data well. We have modelled the counterfactual policy regime (Sys.S) as a Stackelberg game in which a social planner chooses college-specific admissions policies and students make enrollment decisions, choose course portfolios in the trial period, learn about their fits and then choose their majors. We have shown how the distribution of student educational outcomes changes and provided bounds on potential welfare gains from adopting the new system.

Although our empirical application is based on the case of Chile, our framework can be easily adapted to other countries with similar admissions systems. A natural extension to our paper, given data on student performance in college and/or market returns to partial college training, is to model human capital production as a cumulative process and to measure achievement at each stage of one's college life. This extension would allow for a sharper prediction of the impacts on student welfare when the admissions system changes. With data on in-college performance, it is also feasible to model learning as a gradual process, and to allow for additional in-college uncertainties. Another extension would introduce more heterogeneity across colleges besides their student quality and course requirements. A more comprehensive model would allow the social planner to choose college investment together with admissions policies, which would require data on college investment.

One important question arises naturally from our findings: what explains the prevalence of Sys.J? Except in the most pessimistic case, our results reveal that a switch from Sys.J to Sys.S would improve average student welfare. If the goal were to maximize the overall student welfare, the fact that countries like Chile have not switched to Sys.S might be explained by some switching costs, such as increases in college operational costs. A comparison of average student welfare levels suggests that the switching cost would need to be at least as high as 2,760 USD per student for Chile not to make the switch.[74] However, countries may use other criteria in their choices of admissions systems. For example, we also find that Sys.J better serves advantaged students although at the cost of the others.

References

Altonji J. (1993) "The Demand for and Return to Education When Education Outcomes Are Uncertain," *Journal of Labor Economics,* 11(1): 48–83.

[74]See welfare levels in the first and last columns of Table 11.

Altonji J., E. Blom and C. Meghir (2011) "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, 4: 185-223.

Acemoglu, D. (2002) “Technical Change, Inequality, and the Labor Market,” *Journal of Economic Literature*, 40(1): 7–72.

Arcidiacono P. (2004) "Ability Sorting and the Returns to College Major," *Journal of Econometrics,* 121(1–2): 343–375.

Arcidiacono, P., and S. Nicholson (2005): “Peer Effects in Medical Schools,” *Journal of Public Economics,* 89: 327–350.

Betts, J., and D. Morell (1999) “The Determinants of Undergraduate GPA: The Relative Importance of Family Background, High School Resources, and Peer Group Effects,” *Journal of Human Resources,* 107: 797–817.

Blume, L., W. Brock, S. Durlauf and Y. Ioannides (2011) "Identification of Social Interactions," *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin and M. Jackson, North-Holland, Vol 1B: 853-964.

Brock, W., and S. Durlauf (2001) “Discrete Choice with Social Interactions,” *Review of Economic Studies,* 68: 235–260.

Card, D. (2009) "Immigration and Inequality," *American Economic Review: Papers & Proceedings*, 99 (2): 1–21.

Cunha, F., J. Heckman and S. Navarro (2005) “Separating Uncertainty from Heterogeneity in Life Cycle Earnings." *Oxford Economic Papers*, 57(2): 191-261.

Dale, S. and A. Krueger. (2002) "Estimating The Payoff To Attending A More Selective College: An Application Of Selection On Observables And Unobservables," *The Quarterly Journal of Economics*, 117(4): 1491-1527.

Epple, D., R. Romano and H. Sieg. (2006) "Admission, Tuition, And Financial Aid Policies in The Market For Higher Education." *Econometrica,* 74(4): 885-928.

Fu, C. (2014) "Equilibrium Tuition, Applications, Admissions and Enrollment in the College Market." *Journal of Political Economy*, 122(2): 225-281.

Goldin, C., and L. Katz. (2008) "The Race between Education and Technology," Cambridge, MA: Harvard University Press.

Hastings, J., C. Neilson and S. Zimmerman. (2013) "Are Some Degrees Worth More Than Others? Evidence from College Admission Cutoffs in Chile." NBER working paper.

Kinsler, J. and R. Pavan (2014) "The Specificity of General Human Capital: Evidence from College Major Choice." *Journal of Labor Economics*, forthcoming.

Malamud, O. (2010) "Breadth v.s. Depth: The Timing of Specialization in Higher Education." *Labour* 24(4): 359-390.

Malamud, O. (2011) "Discovering One's Talent: Learning from Academic Specialization." *Industrial and Labor Relations Review*. 64(2): 375-405.

Manski, C. (1993) "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60: 531–542.

Manski, C. and D. McFadden (1981) "Alternative Estimators and Sample Designs for Discrete Choice Analysis." in *Structural Analysis of Discrete Data and Econometric Applications*, ed. by Manski, C. and D. McFadden. Cambridge, MA: MIT Press, 2-50.

Moffitt, R. (2001) "Policy Interventions, Low-Level Equilibria, and Social Interactions," in *Social Dynamics*, ed. by S. Durlauf and P. Young. Cambridge, MA: MIT Press, 45–82.

Ottaviano, G., and G. Peri. (2008) "Immigration and National Wages: Clarifying the Theory and the Empirics," NBER Working Paper No. 14188.

Sacerdote, B. (2001) "Peer Effects with Random Assignment: Results from Dartmouth Roommates," *Quarterly Journal of Economics*, 116: 681–704.

Silos, P. and E. Smith (2012) "Human Capital Portfolios," Working paper, Federal Reserve Bank of Atlanta.

Stange, K. (2012) "An Empirical Examination of the Option Value of College Enrollment." *American Economic Journal: Applied Economics*, 4(1): 49–84.

Stinebrickner, T. and R. Stinebrickner (2011) "Math or Science? Using Longitudinal Expectations Data to Examine the Process of Choosing a College Major." NBER Working Paper No. 16869

Wiswall, M. and B. Zafar (2014) "Determinants of College Major Choice: Identification using an Information Experiment," forthcoming *Review of Economic Studies.*

Zafar B. (2009) "College Major Choice and The Gender Gap," Staff Report 364, Federal Reserve Bank, New York, NY

Zimmerman, D. (2003) "Peer Effects in Higher Education: Evidence from a Natural Experiment," *The Review of Economics and Statistics*, 85: 9–23.

Appendix A

A1. A student's idiosyncratic taste consists of three independent parts, such that

$$\epsilon_{jmi} = \varepsilon_{ji} + \varepsilon_{mi} + \varepsilon_{jmi}.$$

$\varepsilon_{ji}\tilde{}\, i.i.d.N(\overline{v}_j, \sigma^2_{\text{col}})$ is one's taste for college j, with $\overline{v}_j$ being the consumption value of college j for an average student. $\varepsilon_{mi}\tilde{}\, N(0, \sigma^2_{\text{major}_m})$ is the taste for major m. $\varepsilon_{jm}\tilde{}\, i.i.d.N(0, \sigma^2_{\text{prog}})$ is one's taste for the specific program.

A2 Adjustment

A2.1 Adjusted Value Functions

The first period in college lasts two years for all majors. Letting the total length of major m be l_m, the adjusted second-period value function is given by

$u_{jm}(x, \epsilon, \widetilde{\eta}_{jm}|A_{jm}) =$

$$\max\left\{\left(\begin{array}{l} \sum_{\tau'=3}^{l_m} \beta^{\tau'-3}(v_{jm}(x,\epsilon,A_{jm}) - C_{jm}(x)) + \\ \sum_{\tau'=l_m+1}^{T} \beta^{\tau'-3}\left[\begin{array}{c} E_\zeta\left(w_m(\tau - l_m - 1, x, A_{jm}, \eta, \rho_{jm}, \zeta)\right) \\ +v_{jm}(x,\epsilon,A_{jm}) \end{array}\right] \end{array}\right), V^d_{jm}(x)\right\}.$$

The adjusted first-period value function is given by

$$\begin{array}{rcl} U(x,\epsilon|a^*,A) & = & \max\left\{\max\limits_{(j,m)}\left\{\begin{array}{c} \beta^2 E_\eta(u_{jm}(x,\epsilon,\widetilde{\eta}_{jm}|A_{jm})) \\ +\sum_{\tau'=1}^{2}\beta^{\tau'-1}(v_{jm}(x,\epsilon,A_{jm}) - C_{jm}(x) \end{array}\right\}, V^0(x)\right\} \\ & & s.t.\ E_\eta(u_{jm}(x,\epsilon,\widetilde{\eta}_{jm}|A_{jm})) = -\infty \text{ if } a_m < a^*_{jm}. \end{array}$$

A2.2 Empirical Definitions of ω, a^* and Retention Rates

1) Programs aggregated in major m have similar weights ω_m. In case of discrepancy, we use the enrollment-weighted average of $\{\omega_{ml}\}_l$ across these programs.
2) For the cutoff a^*_{jm}, we first calculate the adjusted cutoffs using weights defined in 1) and then set a^*_{jm} to be the lowest cutoff among all programs within the (j,m) group.
3) The retention rate in (j,m) is the ratio between the total number of students staying in (j,m) and the total first-year enrollment in (j,m).

A3 Estimation and Equilibrium-Searching Algorithm

Without analytical solutions to the student problem, we integrate out their unobserved tastes numerically: for every student x, draw R sets of taste vectors ϵ. The estimation involves an outer loop searching over the parameter space and an inner loop searching for equilibria. The algorithm for the inner loop is as follows:
0) For each parameter configuration, set the initial guess of o at the level we observe from the data, which is the realized equilibrium.
1) Given o, solve student problem backwards for every (x,ϵ), and obtain enrollment decision $\left\{\delta^1_{jm}\left(x,\epsilon|a^*,A\right)\right\}_{jm}$.[75]
2) Integrate over (x,ϵ) to calculate the aggregate $\{A_{jm}\}_{jm}$, yielding o^{new}.
3) If $\|o^{new}-o\|<\upsilon$, a small number, end the inner loop. If not, $o=o^{new}$ and go to step 1).

This algorithm uses the fact that all equilibrium objects are observed to deal with potential multiple equilibria: we always start the initial guess of o at the realized equilibrium level and the algorithm should converge to o at the true parameter values, moreover, the realized equilibrium o also serves as part of the moments we target.

Additional Tables

1. Data

[75] Conditional on enrollment in (j,m), the solution to a student's continuation problem follows a cutoff rule on the level of $\sum_n \rho_{jmn}\eta_n \sim N\left(0, \sum_n \rho^2_{jmn}\sigma^2_{\eta_n}\right)$, which yields closed-form expressions for $E(u_{jm}(x,\epsilon,\eta||A_{jm}))$. Details are available upon request.

Table A1.1 Score Weights (ω) and Length of Study

	Weights[a] (%)						Length
	Language	Math	GPA	Social Sc	Science	max(Social Sc., Science)[b]	(years)
Medicine	22	30	25	0	23	0	7
Law	33	19	27	21	0	0	5
Engineering	18	40	27	0	15	0	6
Business	21	36	31	0	0	12	5
Health	23	29	28	0	20	0	5
Science	19	36	30	0	15	0	5
Arts&Social	31	23	28	18	0	0	5
Education	30	25	30	0	0	15	5

[a]Weights used to form the index in admissions decisions, weights on the six components add to 100%.

[b]Business and education majors allow student to use either social science or science scores to form their indices, students use the higher score if they took both tests.

Table A1.2 College-Major-Specific Cutoff Index

	Medicine	Law	Engineering	Business	Health	Science	Arts&Social	Education
Tier 1	716	679	597	609	640	597	578	602
Tier 2	663	546	449	494	520	442	459	468
Tier 3	643	475	444	450	469	438	447	460

The lowest admissible major-specific index across all programs within each tier-major category.

Table A1.3 College-Major-Specific Annual Tuition (1,000 Peso)

	Medicine	Law	Engineering	Business	Health	Science	Arts&Social	Education
Tier 1	4,546	3,606	4,000	3,811	3,085	3,297	3,086	3,012
Tier 2	4,066	2,845	2,869	2,869	2,547	2,121	2,292	1,728
Tier 3	4,229	2,703	2,366	2,366	2,391	2,323	2,032	1,763

The average tuition and fee across all programs within each tier-major category.

Table A1.4 Course Credit Weights (ρ) Averaged over Tiers

%	General Courses[a]					Major-specific[b]
	Science	Social Sci.&Language	Econ&Business	Math&Stats	Medical	
Medicine	6.1	1.3	0	1.7	8.2	82.7
Law	0	5.6	5.9	0	0	88.5
Engineering	8.8	2.3	7.8	16.8	0	64.2
Business	0	5.7	10.9	14.1	0	69.3
Health	4.9	3.9	0	2.3	12.1	76.8
Science	33.9	5.3	0	17.6	0	43.1
Arts&Social	0.6	19.3	2.4	5.8	0	71.9
Education	17.6	19.9	0	11.0	0	51.5

Each row add up to 100%. We calculate the credits in a course category as a percentage of the total credits in each (tier, major), then average over tiers for each major.

[a] Courses required in more than one major

[b] Course required exclusively for the major in question

Table A1.5 Average Test Scores

	Medicine	Law	Engineering	Business	Health	Science	Arts&Social	Education
Tier 1	773	709	719	722	696	687	664	661
Tier 2	723	634	619	605	636	590	598	595
Tier 3	704	575	578	557	588	559	545	564

The average $\frac{Math+language}{2}$ across all students within each tier-major category.

Table A1.6 Wage Regressions

	Peer Ability	Own Ability
Medicine	-0.65	0.48
Law	0.45	1.00
Engineering	0.74	1.44
Business	1.64	1.47
Health	0.73	0.35
Science	1.68	0.97
Arts&Social	0.65	1.04
Education	0.84	0.45

Major-Specific wage regressions.

Other controls are experience, experience^2, gender.

2. Parameter Estimates

We fix the annual discount rate at 0.9.[76] Table A2.1 shows the course-specific contribution of own ability to human capital production, i.e., $\{\gamma_n\}_n$. The left panel shows γ_n's in major-specific courses. The right panel shows γ_n's in general courses, where G. stands for "general."

Table A2.1 Human Capital Production: Own Ability (γ)

Major-Specific Courses			General Courses		
Medicine	0.04	(0.01)	G. Science	0.06	(0.01)
Law	1.53	(0.03)	G. Social Sci., Language	0.66	(0.10)
Engineering	1.77	(0.02)	G. Econ and Business	0.14	(0.02)
Business	1.17	(0.01)	G. Math and Stats	1.36	(0.17)
Health	0.01	(0.004)	G. Medicine+Health	0.10	(0.03)
Science	0.03	(0.01)			
Arts&Social	1.12	(0.02)			
Education	1.69	(0.02)			

Table A2.2 shows how the value of one's outside option varies with one's characteristics.[77] The constant term of the outside value for a student from a low income family is only 57% of that for one from a high income family. Relative to a high school graduate, the outside value faced by a college dropout is about 3% higher.

Table A2.2 Outside Value

Constant (θ_{01})	13131.8	(60.1)
Low Income (θ_{02})	0.57	(0.01)
Language (θ_1)	351.0	(5.6)
Math (θ_2)	330.7	(5.2)
Dropout (κ)	1.03	(0.02)

Table A2.3 shows major-independent parameters that govern one's consumption value: the left panel for college programs and the right panel for majors. Relative to Tier 3 colleges, Tier 2 colleges are more attractive to an average student, while top-tier colleges are less attractive.[78] We have restricted σ_{major_m} to be the same across

[76] Annual discount rates used in other Chilean studies range from 0.8 to 0.96.

[77] We cannot reject the hypothesis that the outside value depends only on math and language scores, therefore, we restrict θ_l for other test scores to be zero.

[78] One possible explanation is that the two top tier colleges are both located in the city of Santiago, where the living expenses are much higher than the rest of Chile.

majors that are science-oriented (engineering, science, health and medicine) and the same across the other four majors.[79] The standard deviations of student tastes suggest substantial heterogeneity in student educational preferences.

Table A2.3 Consumption Value (Major-Independent Parameters)

College Value			Major Value		
Tier 1 ($\overline{v}_1$)	-664.0	(285.9)	a_m^2 (λ_{2m})	0.017	(0.001)
Tier 2 ($\overline{v}_2$)	3820.5	(228.3)			
σ_{col}	1094.0	(151.3)	$\sigma_{\text{major\|science-related majors}}$	4875.5	(34.3)
σ_{prog}	3079.0	(150.5)	$\sigma_{\text{major\|other majors}}$	4497.1	(84.0)

$\overline{v}_3$ is normalized to 0.

Table A2.4 shows major-independent cost parameters. The impact of tuition is larger for low-family-income students than their counterpart. A student's costs increase significantly if her ability is far from her peers.

Table A2.4 College Cost (Major-Independent Parameters)

I(Low Inc)*Tuition (c_1)	4.72	(0.09)
I(Low Inc)*Tuition^2 (c_2)	-0.0004	(0.0001)
$(a_m - A_{jm})^2$ (λ_4)	-4.61	(0.13)

Table A2.5 shows parameters in the wage function, other than the effects of own ability and peer quality. It is worth noting that females earn less than their male counterparts in most majors, which contributes to the lower college enrollment rate among females.

[79] We have also tried more flexible specifications, but we cannot reject the null that distribution of tastes are the same within each of the two broad categories.

Table A2.5 Other Parameters in Log Wage Functions

	Constant		Experience		Experience^2		female	
Medicine	7.98	(0.04)	0.09	(0.01)	-0.002	(0.001)	0.006	(0.002)
Law	-1.93	(0.01)	0.10	(0.03)	-0.007	(0.002)	0.22	(0.01)
Engineering	-4.68	(0.10)	0.10	(0.01)	-0.002	(0.001)	-0.20	(0.02)
Business	-11.62	(0.02)	0.11	(0.01)	-0.003	(0.001)	-0.22	(0.02)
Health	3.54	(0.02)	0.02	(0.002)	-0.0001	(0.001)	-0.29	(0.03)
Science	-8.68	(0.02)	0.05	(0.01)	-0.0007	(0.0001)	-0.24	(0.03)
Arts&Social	-5.15	(0.03)	0.05	(0.02)	-0.0001	(0.0001)	0.15	(0.02)
Education	-7.01	(0.04)	0.07	(0.01)	-0.002	(0.001)	-0.75	(0.03)
Wage Shock (σ_ζ)		0.35		(0.04)				

Appendix B

B1 Illustration: Gender Differences

To explore the importance of gender-specific preferences in explaining different enrollment patterns across genders, we compare the baseline model prediction with a new equilibrium where females have the same preferences as males.[80] Table B1 shows the distribution of enrollees within each gender in the baseline equilibrium and the new equilibrium. When females share the same preferences as males, there no longer exists a major that is obviously dominated by one gender. Some differences between male and female choices still exist. For example, although college enrollment rate among females increases from 26.6% to 27.3% (not shown in the Table), it is still lower than that among males (30.1%) . Moreover, compared with males, females are still more likely to enroll in social science. One reason is that, on average, females have lower test scores than females; and they have comparative advantage in majors that uses language more than math.[81]

[80] The purpose of this simulation is simply to understand the importance of preferences; the simulation ignores potential changes in admission cutoffs.

[81] The average math score for males (females) is 572 (547), and the average language score for males (females) is 557 (553).

Table B1 Female Enrollee Distribution

(%)	Baseline		New	
	Male	Female	Male	Female
Medicine	4.5	3.0	5.1	4.9
Law	3.6	3.1	3.3	6.5
Engineering	47.6	25.1	47.7	44.8
Business	9.8	11.0	10.2	9.5
Health	5.4	17.0	5.3	4.1
Science	9.3	8.9	9.2	10.3
Arts&Social	9.8	11.9	9.5	15.4
Education	10.2	19.4	9.8	4.4

B2 Counterfactual Model Details: Sys.S

B2.1 Student Problem

B2.1.1 Continuation Decision

Given her first-period course-taking choice $M(x, \epsilon, j, A_j)$, a student learn about her fits $\widetilde{\boldsymbol{\eta}}_{jM(\cdot)} \equiv \cup_{m \in M(x,\epsilon,j,A_j)} \widetilde{\eta}_{jm}$. Given $\left(x, \epsilon, \widetilde{\boldsymbol{\eta}}_{jM(\cdot)}\right)$ and $A_j \equiv \{A_{jm}\}_m$, an enrollee in college j chooses one major of interest or drops out:

$$u_j(x, \epsilon, \widetilde{\boldsymbol{\eta}}_{jM(\cdot)} | A_j) =$$
$$\max \left\{ \max_{m \in M_a} \left\{ \begin{array}{c} v_{jm}(x, \epsilon, A_{jm}) - C_{jm}(x) + \\ E \sum_{\tau'=3}^{T} \beta^{\tau'-2} \left(\begin{array}{c} w_m(\tau - 3, x, A_{jm}, \widetilde{\eta}_{jm}, \rho_{jm}, \zeta) \\ + v_{jm}(x, \epsilon, A_{jm}) \end{array} \right) \end{array} \right\}, V_{jm}^{d}(x) \right\}.$$

Let $\delta_{m|j}^{2}\left(x, \epsilon, \widetilde{\boldsymbol{\eta}}_{jM(\cdot)} | A_j\right) = 1$ if an enrollee in j with $\left(x, \epsilon, \widetilde{\boldsymbol{\eta}}_{jM(\cdot)}\right)$ chooses major m.

B2.1.2 Enrollment and Course Choices

We assume that in the first period of college, an enrollee who allocates her time according to $\{\pi_m\}_m$ pays the weighted averaged cost for and derives the weighted averaged consumption value from various majors, with weights given by $\{\pi_m\}_m$. A student chooses the best among colleges she is eligible for and the outside option:

$$
\begin{aligned}
&U\left(x,\epsilon|q\left(a\right),A\right)= \\
&\max\left\{\max_{j}\{\beta E_{\eta}u_{j}(x,\epsilon,\widetilde{\boldsymbol{\eta}}_{jM(\cdot)}|A_{j})+\sum_{m}\pi_{m}(v_{jm}(x,\epsilon,A_{jm})-C_{jm}\left(x\right))\},V^{0}\left(x\right)\right\} \\
&s.t.\ E_{\eta}u_{j}(x,\epsilon,\widetilde{\boldsymbol{\eta}}_{jM(\cdot)}|A_{j})=-\infty\text{ if }\psi_{j}\left(q\left(a\right)\right)=0, \\
&\pi_{m}=\frac{E(\delta_{m|j}^{2}\left(x,\epsilon,\widetilde{\boldsymbol{\eta}}_{jM(\cdot)}|A_{j}\right)=1)}{\sum_{m}E(\delta_{m|j}^{2}\left(x,\epsilon,\widetilde{\boldsymbol{\eta}}_{jM(\cdot)}|A_{j}\right)=1)}.
\end{aligned}
$$

where $q\left(a\right)$ is the planner's admissions rule for a student with ability a, and $\psi_{j}\left(q\left(a\right)\right)=1$ if the student is eligible for college j. The last constraint requires student's choice of $\left\{\pi_{m}\right\}_{m}$ be consistent with her expected second-period major choice. Let $\delta_{j}^{1}\left(x,\epsilon|q\left(a\right)\right)=1$ if the student chooses college j and $\pi_{m}\left(x,\epsilon,j\right)$ be the consistent π_{m} for an enrollee in j with $\left(x,\epsilon\right)$.

B2.2 Planner's Problem

To formalize the constraint on the planner's strategy space, we introduce the following notation. Let $\Xi\equiv\left\{\chi_{1},\chi_{2},\chi_{3},\chi_{4}\right\}=\left\{\left[1,1,1\right],\left[0,1,1\right],\left[0,0,1\right],\left[0,0,0\right]\right\}$, where the j-th component of each χ_{n} represents the admissions to college j, i.e., $\chi_{nj}=1$ if a student is eligible for college j. Denote the planner's admissions policy for student with ability a as $q\left(a\right)$, we restrict the planner's strategy space to be probabilities over Ξ. That is, for all a, $q\left(a\right)\in Q\equiv\Delta\left(\left[1,1,1\right],\left[0,1,1\right],\left[0,0,1\right],\left[0,0,0\right]\right)$, a convex and compact set. The probability that a student is eligible for college j, denoted as $\psi_{j}\left(q\left(a\right)\right)$, is given by $\psi_{j}\left(q\left(a\right)\right)=\sum_{n=1}^{4}q_{n}\left(a\right)\chi_{nj}$.

Consistent with the assumptions on student course taking, we assume that in the first period in college, a student with choice $\left\{\pi_{m}\right\}_{m}$ will take π_{m} slot in major m, and that in the second period in college, she will take one slot in her chosen major and zero slot in other majors. Let $z=\left[y,I\left(female\right)\right]$ be the part of x that is not observable to the planner, the planner's problem reads:

$$
\pi=\max_{\{q(a)\in Q\}}\left\{\int_{a}\widetilde{U}\left(a|q\left(a\right),A\right)f_{a}(a)da\right\}
$$

where $\widetilde{U}\left(a|q\left(a\right),A\right)=\int_{z}\int_{\epsilon}U\left(x,\epsilon|q\left(a\right),A\right)dF_{\epsilon}\left(\epsilon\right)dF_{z}\left(z|a\right)$ is the expected utility of student with ability a, integrating out student characteristics that are unobservable to the planner.[82]

For each a, one can take the first order conditions with respect to $\{q_l(a)\}_{l=1}^{4}$, subject to the constraint that $q(a) \in Q$. Given the nature of this model, the solution is generically at a corner with one of the $q_l(a)$'s being one. Thus, we use the following algorithm to solve the planner's problem. For each student a, calculate the net benefit of each of the four pure strategies ($[1,1,1], [0,1,1], [0,0,1], [0,0,0]$). The (generically unique) strategy that generates the highest net benefit is the optimal admissions policy for this student. Let "$\cdot$" stand for $(q(a), A)$, it can be shown that the net benefit of applying some $q(a)$ to student with ability a is:

$$\begin{aligned}
&f_a(a)\int_z\int_\epsilon U(x,\epsilon|\cdot)\,dF_\epsilon(\epsilon)dF_{z|a}(z) \qquad (12)\\
&+f_a(a)\sum_j \psi_j(\cdot)\delta_j^1(a|\cdot)\sum_m (a_m - A_{jm})\,\pi_m(\cdot)\,b_m\varphi_m A_{jm}^{\varphi_m-1}K_{jm}\\
&-f_a(a)\sum_j \psi_j(\cdot)\delta_j^1(a|\cdot)\sum_m (a_m - A_{jm})\,\pi_m(\cdot)\begin{pmatrix} \lambda_{3m}(1+\sum_{\tau'=1}^{2}\beta^{\tau'-1}\frac{\mu_{jm}^2}{\mu_{jm}^1}) \\ +2\lambda_4\sum_{\tau'=1}^{2}\beta^{\tau'-1}\frac{\mu_{jm}^2}{\mu_{jm}^1}(A_{jm}-A'_{jm})\end{pmatrix}.
\end{aligned}$$

Elements in (12) will be defined in the next paragraph. The first line of (12) is the expected individual net benefit for student a. An individual student has effect on her peer's net benefits because of her effect on peer quality: the second line calculates her effect on her peers' market return; the third line calculates her effect on her peers' effort costs. Peers of student a are those who study in the programs she takes courses in. Student a's effect on her peers is weighted by her course-taking intensity $\pi_m(\cdot)$.

To be more specific, $\delta_j^1(a|\cdot) = \int_z\int_\epsilon \delta_j^1(x,\epsilon|\cdot)dF_\epsilon(\epsilon)dF_{z|a}(z)$ is the probability that a student with ability a matriculates in college j. $\psi_j(\cdot)\delta_j^1(a|\cdot)$ is the probability that student a is enrolled in college j. μ_{jm}^1 is the size of program (j,m) in the first period, where a student $\pi_m(\cdot)$ seat in major m. A_{jm} is the average ability among these students.

$$\begin{aligned}
\mu_{jm}^1 &= \int_a \delta_j^1(a|\cdot)\psi_j(\cdot)\pi_m(\cdot)\,f_a(a)da,\\
A_{jm} &= \frac{\int_a \psi_j(\cdot)\delta_j^1(a|\cdot)\pi_m(\cdot)\,a_m f_a(a)da}{\mu_{jm}^1}.
\end{aligned}$$

[82]Given that test scores are continuous variables, we nonparametrically approximate $F_{z|a}(z)$ by discretizing test scores and calculating the data distribution of z conditional on discretized scores. In particular, we divide math and language test scores each into l narrowly defined ranges and hence generate l^2 bins of test scores. All a's in the same bin share the same $F_{z|a}(z)$.

The second line of (12) relates to market return. b_m is the part of expected lifetime income that is common to all graduates from major m.[83] K_{jm} is the average individual contribution to the total market return among students who take courses in (j,m) :

$$K_{jm} \equiv \frac{\int_a \psi_j(\cdot) k_{jm}(a) f_a(a)da}{\mu^1_{jm}},$$

$$k_{jm}(a) =$$

$$\int_z e^{\alpha_{3m} I(female)} \int_\epsilon \delta^1_j(x,\epsilon|\cdot)\pi_m(\cdot) \int_\eta \delta^2_{m|j}(x,\epsilon,\eta|A_j)\, a_m^{\sum_n \rho_{jmn}\gamma_n} e^{\left(\sum_n \rho_{jmn}\eta_n\right)} dF_\eta(\eta)\, dF_\epsilon(\epsilon) dF_{z|a}(z).$$

Students with higher a_m contribute more to the total market return of their peers. The third line of (12) relates to effort cost. μ^2_{jm} is the size of program (j,m) in the second period. A'_{jm} is the average ability among students enrolled in (j,m) in the second period. Formally,

$$\begin{aligned}
\mu^2_{jm} &= \int_a \delta^1_j(a|\cdot)\psi_j(\cdot)\delta^2_{m|j}(a|\cdot)\, f_a(a)da, \\
A'_{jm} &= \frac{\int_a \psi_j(\cdot)\delta^1_j(a|\cdot)\delta^2_{m|j}(a|\cdot)\, a_m f_a(a)da}{\mu^2_{jm}},
\end{aligned}$$

where $\delta^2_{m|j}(a|\cdot) = \frac{\int_z\int_\epsilon \delta^1_j(x,\epsilon|\cdot)\int_\eta \delta^2_{m|j}(x,\epsilon,\eta)dF_\eta(\eta)dF_\epsilon(\epsilon)dF_{z|a}(z)}{\delta^1_j(a|\cdot)}$ is the probability that student a will take a full slot in (j,m) in the second period conditional on enrollment in j.

B2.3 Equilibrium

Definition 2 *An equilibrium in this new system consists of a set of student enrollment and continuation strategies* $\left\{\delta^1_j(x,\epsilon|q(a),A), \left\{\delta^2_{m|j}(x,\epsilon,\widetilde{\boldsymbol{\eta}}_{jM(\cdot)}|A_j)\right\}_m\right\}_j$ *, a set of admissions policies* $\{q^*(a)\}$ *, and a set of program-specific vectors*

[83] $b_m = E\left(e^\zeta\right)\sum_{\tau'=3}^T \beta^{\tau'-1} e^{\left(\alpha_{0m}+\alpha_{1m}(\tau'-3)-\alpha_{2m}(\tau'-3)^2\right)}$, so that the expected major-m market value of student with ability a can be written as

$$\begin{aligned}
& b_m \int_z e^{\alpha_{3m} I(female)} \int_\epsilon \delta^1_j(x,\epsilon|\cdot) \int_\eta \delta^2_{m|j}(x,\epsilon,\eta|A_j)\, h(a_m, A_{jm}, \eta)\, dF_\eta(\eta)\, dF_\epsilon(\epsilon) dF_{z|a}(z) \\
= \; & b_m \int_z e^{\alpha_{3m} I(female)} \int_\epsilon \delta^1_j(x,\epsilon|\cdot) \int_\eta \delta^2_{m|j}(x,\epsilon,\eta|A_j)\, a_m^{\sum_n \rho_{jmn}\gamma_n} A_{jm}^{\varphi_m} e^{\sum_n \rho_{jmn}\eta_m} dF_\eta(\eta)\, dF_\epsilon(\epsilon) dF_{z|a}(z) \\
= \; & b_m A_{jm}^{\varphi_m} \int_z e^{\alpha_{3m} I(female)} \int_\epsilon \delta^1_j(x,\epsilon|\cdot) \int_\eta \delta^2_{m|j}(x,\epsilon,\eta|A_j)\, a_m^{\sum_n \rho_{jmn}\gamma_n} e^{\sum_n \rho_{jmn}\eta_m} dF_\eta(\eta)\, dF_\epsilon(\epsilon) dF_{z|a}(z).
\end{aligned}$$

$\{\Omega_{jm}\}_{jm} \equiv \left\{\mu^1_{jm}, \mu^2_{jm}, K_{jm}, A_{jm}, A'_{jm}\right\}_{jm}$, *such that*
(a) $\left\{\delta^2_{m|j}(x, \epsilon, \widetilde{\boldsymbol{\eta}}_{jM(\cdot)}|A_j)\right\}_m$ *is an optimal choice of major for every* $(x, \epsilon, \widetilde{\boldsymbol{\eta}}_{jM(\cdot)})$ *and* A_j;
(b) $\left\{\delta^1_j(x, \epsilon|q(a), A)\right\}_j$ *is an optimal enrollment decision for every* (x, ϵ), *for all* $q(a)$ *and* A;
(c) $q^*(a)$ *is an optimal admissions policy for every* a;
(d) $\{\Omega_{jm}\}$ *is consistent with* $\{q^*(a)\}$ *and student decisions.*

B2.3.1 Equilibrium-Searching Algorithm:

We use the same random taste vectors ϵ for each student as we did for the estimation. In the new model, student continuation problem does not have analytical solutions, so we also draw K sets of random efficiency vectors η. Finding a local equilibrium can be viewed as a classical fixed-point problem, $\Gamma : O \rightrightarrows O$, where $O = \left([0,1] \times [0,1] \times \left[0, \overline{A}\right] \times \left[0, \overline{A}\right] \times \left[0, \overline{K}\right]\right)^{JM}$, $o = \Omega_{jm} \in O$. Such a mapping exists, based on this mapping, we design the following algorithm to compute equilibria numerically.

0) Guess $o = \{\Omega_{jm}\}_{jm} \equiv \left\{\mu^1_{jm}, \mu^2_{jm}, K_{jm}, A_{jm}, A'_{jm}\right\}_{jm}$.
1) Given o, for every (x, ϵ) and every pure strategy $q(a)$, solve the student problem backwards, where the continuation decision involves numerical integration over efficiency shocks η. Obtain $\delta^2_{m|j}(x, \epsilon|q(a))$ and $\delta^1_j(x, \epsilon|q(a))$.
2) Integrate over (ϵ, z) to obtain $\delta^2_{m|j}(a|q(a))$, $\delta^1_j(a|q(a))$ and $\widetilde{U}(a|q(a), A)$.
3) Compute the net benefit of each $q(a)$, and pick the best $q(a)$ and the associated student strategies. Do this for all students, yielding o^{new}.
4) If $\|o^{new} - o\| < \upsilon$, where υ is a small number, stop. Otherwise, set $o = o^{new}$ and go to step 1).

B2.3.2 Global Optimality

After finding the local equilibrium, we verify ex post that the planner's decisions satisfy global optimality. Since it is infeasible to check all possible deviations, we use the following algorithm to check global optimality.[84] Given a local equilibrium $o = \left\{\mu^1_{jm}, \mu^2_{jm}, K_{jm}, A_{jm}, A'_{jm}\right\}_{jm}$, we perturb o by changing its components for a random program (j, m) and search for a new equilibrium as described in B2.3.1. If the algorithm converges to a new equilibrium with higher welfare, global optimality is

[84] Epple, Romano and Sieg (2006) use a similar method to verify global optimality ex post.

violated. After a substantial random perturbations with different magnitudes, we have not found such a case. This suggests that our local equilibrium is a true equilibrium.

B2.4 Endogenous Extension Time $ext\,(\pi)$

To reflect the fact that colleges are normally not fully flexible and only provide two graduation seasons per year, we assume that the extra time in college is measured in mulitiples of semesters (0.5 years) and that $ext(\pi)$ is a step function. In particular, if a student i divide her time across majors according to $\pi_i = \{\pi_{im}\}_m$ and if her intended final major is m, then her extra years in college is given by

$$ext(\pi_i) = \begin{cases} 0 \text{ if } \pi_{im} \geq 7/8 \\ 0.5 \text{ if } \pi_{im} \in [5/8, 7/8) \\ 1 \text{ if } \pi_{im} \in [3/8, 5/8) \\ 1.5 \text{ if } \pi_{im} \in [1/8, 3/8) \\ 2 \text{ if } \pi_{im} < 1/8 \end{cases} . \tag{13}$$

Given that the first period is 2 years or 4 semesters, $\frac{1}{8}$ of the first period is equivalent to half a semester. Therefore, (13) specifies that if a student has spent more than 3.5-semester-worth time in major m in the first period, she can graduate on time (the first line); otherwise if she has spent over 2.5-semester-worth of time in major m, she needs to spend one more semester (0.5 years) in college (the second line) etc. We have also tried other cutoffs (e.g., instead of half a semester, we used multiples of $\frac{3}{4}$ of a semester or multiples of one semester to form the cutoffs), the results are qualitatively similar.

B3 A Closer Look at Sys.S: Gainers and Losers

To illustrate who are more likely to gain/lose, we generate an indicator variable that reflects whether the change in a student welfare is positive, zero or negative. Then, we run an ordered logistic regression of this indicator on student observable characteristics. Table B3 shows the regression results. Females and students from low income families are more likely to be gainers than their counterparts. Students with higher math scores are more likely to gain, while neither language score nor high school GPA has significant effects. A welfare loss is more likely for students with higher score in their track-specific subjects (science or social science) and for those with a larger gap between language score and math score. In other words, when a student has a clear comparative advantage, the cost of delayed specialization is likely to outweigh its benefit.

Table B2 Welfare Gain and Student Characteristics

	Female	Low Income	$\frac{Language}{1000}$	$\frac{Math}{1000}$	$\frac{HSGPA}{1000}$	$\frac{Subject}{1000}$	$\frac{(language-math)^2}{1000}$
Coefficient	0.42**	9.62**	-1.83	9.82**	0.63	-2.72*	-0.025**
Std. Dev.	0.10	0.43	0.96	0.83	0.72	0.86	0.0058

Ordered logistic regression, dependent variable in order: positive/zero/negative welfare change.

Num of Obs: 10,000

* significant at 5% level, ** significant at 1% level.

B4 Proof of existence in a simplified (baseline) model.

Assume there are two programs $m \in \{1, 2\}$ and a continuum of students with ability $a \in \left[0, \overline{A}\right]^2$ that are eligible for both programs. Let the average ability in program j be A_m.[85] The utility of the outside option is normalized to 0. The utility of attending program 1 is $v_1(a, A_1)$ for all who have ability a, and $v_2(a, A_2) - \epsilon$, where ϵ is i.i.d. idiosyncratic taste, a continuous random variable.

Definition 3 *A sorting equilibrium consists of a set of student enrollment strategies* $\{\delta_m(a, \epsilon|, \cdot)\}_m$ *, and the vector of peer quality* $A = [A_1, A_2]$ *, such that*
(a) $\{\delta_m(a, \epsilon|A)\}_m$ *is an optimal enrollment decision for every* (a, ϵ) *;*
(c) A is consistent with individual decisions such that, for $m \in \{1, 2\}$ *,*

$$A_m = \frac{\int_a \int_\epsilon \delta_m(a, \epsilon|A) a_m dF_\epsilon(\epsilon) dF_x(x)}{\int_a \int_\epsilon \delta_m(a, \epsilon|A) dF_\epsilon(\epsilon) dF_x(x)}. \quad (14)$$

Proposition 1 *A sorting equilibrium exists.*

Proof. The model can be viewed as a mapping

$$\Gamma : O \rightrightarrows O,$$

where $O = \left[0, \overline{A}\right]^2$, $o = [A_1, A_2]$ .. The following shows that the conditions required by Brouwer are satisfied and hence a fixed point exists.

[85]It can be shown that conditional on enrollment in a program, the solution to a student's continuation problem follows a cutoff rule on the level of efficiency shock η_m, which yields closed-form expressions for $E_{\eta_m}(u(a, \epsilon, \eta_m|A_m))$. As such, $v_m(\cdot)$ can be viewed as the net expected utility of enrollment, i.e., the difference between $E_{\eta_m}(u(a, \epsilon, \eta_m|A_m))$ and the cost $C_m(a_m, A_m)$, both are continuous functions. Details are available upon request.

1) The domain of the mapping $O = \left[0, \overline{A}\right]^2$ is compact and convex.

2) Generically, each student has a unique optimal enrollment decision. In particular, let $\epsilon^* (a, A) \equiv v_2 (a, A_2) - \max \{0, v_1 (a, A_1)\}$

$$\delta (a, \epsilon | A) = \left\{ \begin{array}{l} [0,1] \text{ if } \epsilon < \epsilon^* (a, A) \\ [1,0] \text{ if } v_1 (a, A_1) > 0 \text{ and } \epsilon \geq \epsilon^* (a, A) \\ [0,0] \text{ if } v_1 (a, A_1) \leq 0 \text{ and } \epsilon \geq \epsilon^* (a, A) \end{array} \right\}.$$

Given that both $v_a (a, A_a)$ are continuous functions of (a, A), so are $\max \{0, v_1 (a, A_1)\}$ and $\epsilon^* (a, A)$.

3) Given the result from 2), the population of students with different (a, ϵ) can be aggregated continuously into the total enrollment in program m via $\int_a \int_\epsilon \delta_m(a, \epsilon|A) dF_\epsilon (\epsilon) dF_x(x)$ and the total ability in m via $\int_a \int_\epsilon \delta_m(a, \epsilon|A) a_m dF_\epsilon (\epsilon) dF_x(x)$, hence the right hand side of (14), being a ratio of two continuous functions, is continuous in A. That is, the mapping Γ is continuous.

4) "Every continuous function from a convex compact subset K of a Euclidean space to K itself has a fixed point." (Brouwer's fixed-point theorem) ■

In the full model, where there are more than two programs and the taste shock is a vector, there will be cutoff hyperplanes. It is cumbersome to show, but the logic of the proof above applies.

Model Fit

Table B3 Enrollment (Low Income) (%)

	Data	Model
Tier 1	2.3	2.5
Tier 2	12.6	12.0
Tier 3	9.7	9.5

Enrollment among students with low family income.

Table B4 Enrollee Distribution Across Majors (Low Income) (%)

	Data	Model
Medicine	1.7	1.5
Law	3.4	3.1
Engineering	35.1	34.8
Business	10.0	10.1
Health	12.2	11.2
Science	8.2	8.9
Arts&Social	11.0	12.3
Education	18.5	18.5

Distribution across majors among enrollees with low family income.

Table B5 Correlation (starting wage, own ability)

	Tier 1		Tier 2		Tier 3	
	Data	Model*	Data	Model	Data	Model
Medicine	0.01	0.01	-0.01	-0.01	0.01	0.02
Law	-0.07	-0.01	0.08	-0.004	-0.06	-0.01
Engineering	0.09	0.04	0.13	0.05	0.09	0.05
Business	0.14	0.04	0.14	0.04	0.02	0.04
Health	-0.05	-0.01	0.08	0.01	0.08	0.02
Science	0.002	0.003	0.15	0.02	0.21	0.05
Arts&Social	0.19	0.05	0.05	0.02	0.09	0.05
Education	0.12	0.04	0.02	0.07	0.11	0.06

*Averaged corr(w_{jm}, a_m) over 100 simulated cases, each using an i.i.d. set of random draws to simulate the equilibrium.

Table B6 Mean Test Scores Among Outsiders

	Data	Model
Math	533	534
Language	532	534
HS GPA	542	542
Max(Science, Soc Science)	531	533

Mean test scores among students who chose the outside option.

Appendix C. Other Examples of Sys.J[86]

C1 China (Mainland)

1. High School Track: Students choose either science or social science track in the second year of high school and receive more advanced training corresponding to the track of choice.

2. College Admissions: At the end of high school, college-bounding students take national college entrance exams, including three mandatory exams in math, Chinese and English, and track-specific exams. A weighted average of the national exam scores forms an index of the student, used as the sole criterion for admissions. College admissions are college-major specific: a student is eligible for a college-major pair if her index is above the program's cutoff.[87]

3. Transfer Policies: Transfers across majors are either near impossible (e.g., between a social science major and a science major) or very rare (e.g., between similar majors).[88]

C2 Japan

1. High School Track: similar to the case in C1.

2. College Admissions: Students applying to national or other public universities take two entrance exams. The first is a nationally administered uniform achievement test, which includes math, Japanese, English and specific subject exams. Different college programs require students to take different subject exams. The second exam is administered by the university that the student hopes to enter. A weighted average of scores in various subjects from the national test forms the first component of the admissions index; a weighted average of university-administered exam scores forms the other. The final index is a weighted average of these two components. College admissions are college-major specific in most public universities, except for the University of Tokyo, which uses category-specific admissions (there are six categories, each consists of a number of majors).

[86] Major Sources of Information: 1. "Survey of Higher Education System" (2004), OECD Higher Education Programme, 2. OECD Reviews of Tertiary Education (by country), 3. Department of Education (by country), 4. Websites of major public colleges in each country.

[87] The cutoffs may be different based on the student's home province.

[88] In 2001, Peking University started a small and very selective experiment program which admits students to two broad areas (social science or science) according to their high school track. Students are free to choose majors within their areas in upper college years.

3. Transfer Policies:

1) University of Tokyo: Students choose one major within the broad category in their sophomore year. After that, a student can transfer to a different major within her current category but only with special permission and she has to spend one extra year in college, besides meeting the grade requirement of the intended major. Transfer across categories is rarely allowed.

2) Other public universities: Changing majors is normally possible only with special permission at the end of the sophomore year, and it may require much make-up or an extra year in college.

C3 Spain

1. High School Track: similar to the case in C1, but with three tracks to choose from: arts, sciences and technology, and humanities and social sciences.

2. College Admissions: All public colleges use the same admissions procedure. College-bounding students take the nation-wide Prueba de Acceso a la Universidad (PAU) exams, which consist of both mandatory exams and track-specific exams. Admissions are college-major specific, and the admissions criterion is a weighted average of student high school GPA and the PAU exam scores.

3. Transfer Policies: Transfers across majors require that the student have accumulated a minimum credit in the previous program that is recognized by the intended program, where the recognition depends on the similarity of the contents taught in the two programs. Transfers across similar majors can happen, although not common, in which cases, the student usually has to spend one extra year in college. Transfers across very different majors are rarely allowed.

C4 Turkey

1. High School Track: Students in regular high schools choose, in their second year, one of four tracks: Turkish language–Mathematics, Science, Social Sciences, and Foreign Languages. In Science High Schools only the Science tracks are offered.

2. College Admissions: Within the Turkish education system, the only way to enter a university is through the Higher Education Examination-Undergraduate Placement Examination (YGS-LYS). Students take the Transition to Higher Education Examination (YGS) in April. Those who pass the YGS are then entitled to take the Undergraduate Placement Examination (LYS) in June, in which students have to answer 160 questions(Turkish language(40), math(40), philosophy(8), geography(12), history(15), reli-

gion culture and morality knowledge(5), biology(13), physics(14) and chemistry(13)) in 160 minutes. Only these students are able to apply for degree programs. Admissions are college-major specific and students are placed in courses according to their weighted scores in YGS-LYS.

3. Transfer Policies: Most universities require a student meet strict course and GPA requirement and provide faculty reference in order to transfer majors. In a few universities, the transfer policies are more flexible. However, transfers across very different majors are near infeasible and transfers across similar majors are uncommon as well.

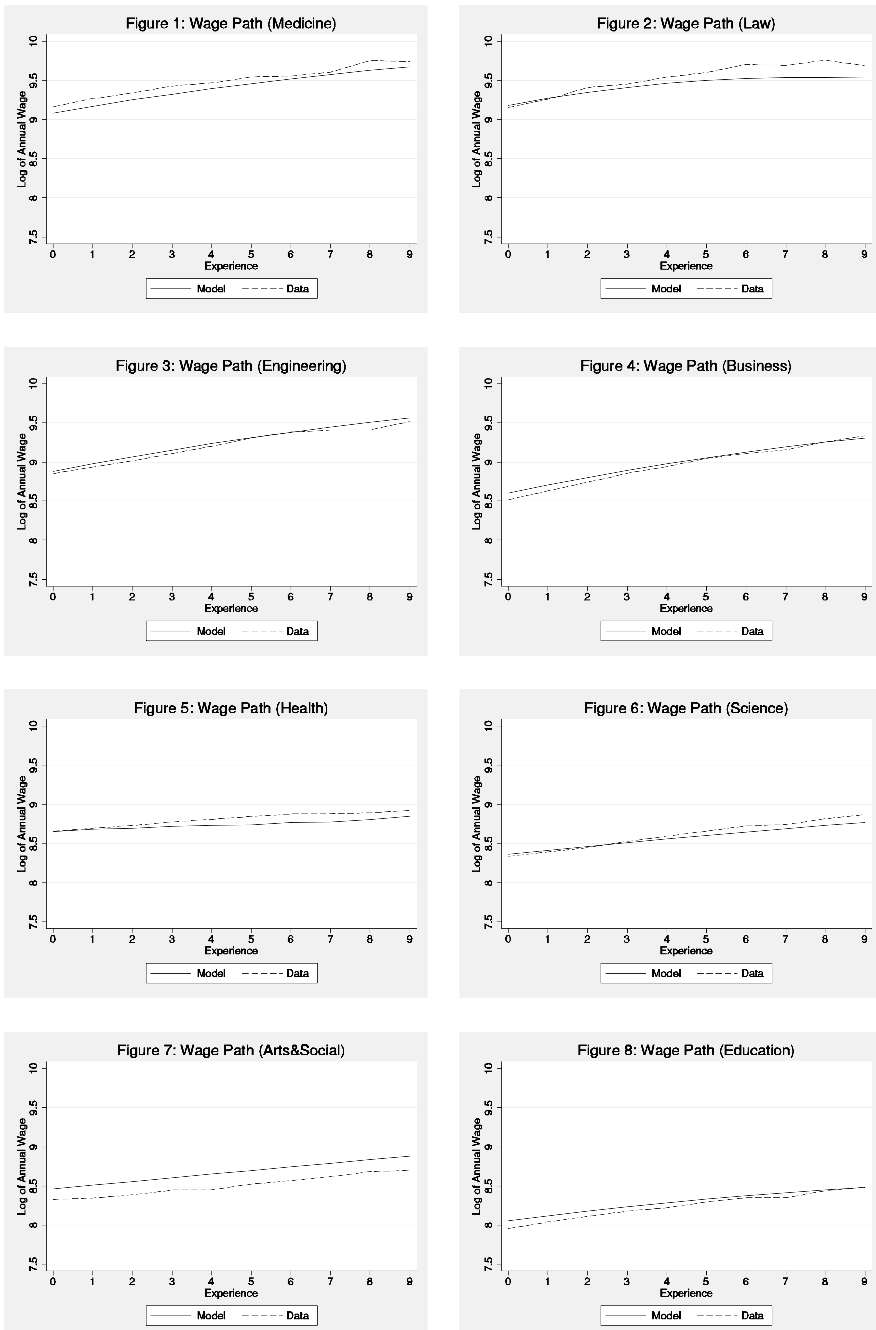

Figure 1: Wage Profile by Experience