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

Paola Bordon   Chao Fu†

August 16, 2013

Abstract

Many countries use college-major-specific admissions policies that require a student to choose a college-major pair jointly. Motivated by potential 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. We compare outcomes and welfare under the two different regimes.

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

*We thank Fumihiko Suga for excellent research assistance. We thank 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 for their insightful discussions. We thank 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 for helpful comments. All errors are ours. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

†Bordon and Fu: Department of Economics, University of Wisconsin-Madison. Please send correspondence to Chao Fu, 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 years 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 in Asia, Europe and Latin America 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 upon enrollment. We label the first system where students choose majors after enrollment by Sys.S (for sequential), and the second system where students have to make a joint college-major choice by Sys.J (for joint).

Which system is more efficient 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 more efficient: 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 serious mismatch problems. Efficiency comparisons across these two admissions systems depend critically on the degree of uncertainty faced by students, the relative importance of peer effects, and student sorting behavior that determines equilibrium peer quality. Simple cross-system comparisons are unlikely to be informative because of the potential 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 comparison and providing necessary information for policy makers before implementing 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

[1] With the exception of Quebec province.

to other countries with similar admissions systems.

In the model, students differ in their (multi-dimensional) abilities and educational preferences, and they face uncertainty about their suitability to 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 first set of counterfactual policy experiments (Sys.S), a planner chooses optimal college-specific, rather than college-major-specific, admissions policies; students make enrollment decisions and postpone their choices of majors until after they learn about their fits to various majors. Taking into account the externality of 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 overall college retention rate from 75% in the baseline to 90% in the counterfactual.

Second, in contrast to Sys.J, where peer students are from the same major upon college enrollment, Sys.S features a much broader student body in first-period classes. While students differ in their comparative advantages, some students have advantages over others 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 "elite" students in their first-period classes. The overall efficiency depends on, among other factors, which of the 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. Combining this fact with the improvement in student-major match quality, we find that the average productivity of college graduates improves in all majors when Sys.S is adopted.

Finally, as students spend time trying out different majors, their specialized training is delayed. Welfare comparisons vary with how costly this delay is. Average student welfare will increase by 5%, 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 a student has to make up for this loss by extending her college life accordingly, a 0.9% loss in mean welfare will result.

As there are both pro's and con's to each system, our second counterfactual experiment involves a hybrid policy that combines the merits of Sys.S and Sys.J: college enrollees are allowed to choose either to specialize upon enrollment or to postpone the choice of major until after they learn about their fits to various majors. With this extra flexibility, the hybrid policy leads to even higher welfare gains than Sys.S does: the higher the cost of postponing specialization, the greater the advantage of the hybrid policy.

Given that under the same admissions framework of Sys.J, some countries have more flexible transfer policies than others, we also conduct an experiment to explore the gains from such flexibility. We find that a more flexible transfer policy always improves welfare but the improvement is milder than that from a switch to Sys.S as long as the cost of delaying specialization is not too high.

Our paper is closely related to studies that treat education as a sequential choice made under uncertainty and emphasize the multi-dimensionality of human capital.[2] 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., where students learn about their abilities via test scores in college before settling down to their majors. As in our paper, he allows for peer effects.[3] Focusing on individual decisions, he treats peer quality as exogenous.[4]

While this literature has focused on individual decision problems, our goal is to

[2]Examples of theoretical papers include Manski (1989) and Comay, Melnick and Pollachek (1973).

[3]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 (1998) have mixed findings.

[4]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.

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 (2013). Both papers study college enrollment in a decentralized market, where colleges compete for better students.[5] 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 (2013).[6] Instead, we emphasize some other 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 on the comparison across different admissions systems are relatively sparse. Ofer Malamud has a series of papers that compare the labor market consequences between the English (Sys.J) and Scottish (Sys.S) systems.[7] Malamud (2010) finds that the 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 sufficiently large to outweigh the greater loss in skills from specializing early. With the caveat that students in two countries may differ in unobservable ways, his findings contribute to our understanding of the relative merits of the two systems. Our paper aims at pushing the frontier toward comparing the relative efficiency of alternative systems for *the same population of students.*

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

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

[6] College-provided financial aid and scholarships are rare in Chile.

[7] "English students usually follow a narrow curriculum that focuses on the main field and allows for little exposure to other fields. Indeed, most universities in England require students who switch fields of study to start university anew (though several do allow for some limited switching across related fields)." (Malamud (2010)).

lays out the model. Section 4 describes the data. Section 5 describes the estimation followed by the empirical results. Section 7 conducts counterfactual policy experiments. The last section concludes the paper. The appendices in the paper and online 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 want to pursue a college degree 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 accounts for over 60% of all Chilean students enrolled in the higher education system.[8] 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.

The traditional universities employ a single admission process: the University Selection Test (PSU), which is very 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.[9] 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. A student is admitted to a specific college-major pair if her index

[8] Enrollments in professional institutes and technical formation centers account for 25.7% and 13.7% respectively.

[9] In 2011, the fee was $23,500$ pesos (1 USD is about 484 Chilean pesos).

is above the cutoff index required by that program. That is, college admissions are college-major specific. A student must choose a college-major pair jointly.

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.

Transfers across programs are rare in Chile. 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."[10] The same report also notes that the highly inflexible curriculum design further limits the mobility between programs.[11] If a student dropped out in order to re-apply to other programs in traditional universities, she must re-take the PSU test.[12]

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. The online Appendix B4 provides further descriptions of the systems in these other countries.

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

[11] "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.

[12] 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 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 first period in college, she learns about her fit to her major and decides whether or not to continue her studies. To smooth reading, we leave some details of the model to the end of this section and some functional form assumptions in the appendix.

3.1 Primitives

There is a continuum of students faced with 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.

3.1.1 Student Characteristics

Students differ in gender (g), family income group (y), abilities (a) and academic interests. In particular, $a = \{a_m\}_{m=1}^M$ is a vector of major-specific (pre-college) abilities. Denote student characteristics that are observable to the researcher by the vector $x \equiv [a, y, g]$, and its distribution by $F_x(\cdot)$.

3.1.2 Skills and Wages

Skill attainments in college depends on a student's major-specific ability (a_m), peer quality (A_{jm}), and how efficient/suitable she is for the major.[13] In particular, A_{jm} is the average major-m ability of enrollees in (j, m).[14] A student faces uncertainty in making enrollment decisions because major-specific efficiency is revealed to her only *after* she takes courses in that major. Denote one's major-specific efficiency as $\{\eta_m\}_m \,\tilde{}\, i.i.d. F_\eta(\cdot)$.

[13] 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.

[14] 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.

The human capital production function is[15]

$$h_m\left(a_m, \eta_m, A_{jm}\right) = a_m^{\gamma_{1m}} A_{jm}^{\gamma_{2m}} \eta_m.$$

Wages are major-specific stochastic functions of one's human capital (hence of a_m, η_m, A_{jm}), work experience (τ) and one's observable characteristics besides ability, where the randomness comes from a transitory wage shock ζ_τ. Denote the wage rate for a graduate from program (j, m) by $w_m\left(\tau, x, \eta_m, A_{jm}, \zeta_\tau\right)$.[16]

3.1.3 Consumption Values and Costs

The per-period net consumption value of attending (j, m) is

$$v_{jm}(x, \epsilon, A_{jm}) = v_m(x, \epsilon_{1m}) + \epsilon_{2jm} - C_{jm}(x, A_{jm}), \tag{1}$$

where $v_m(x, \epsilon_{1m})$ is the consumption value of major m and $C_{jm}(x, A_{jm})$ is the cost of attending program (j, m). ϵ_{1m} is one's taste for major m and ϵ_{2jm} that for program (j, m). Let $\epsilon_1 = \{\epsilon_{1m}\}_m$, $\epsilon_2 = \{\epsilon_{2jm}\}_{jm}$ and $F_\epsilon(\cdot)$ be the joint distribution of $\epsilon = [\epsilon_1, \epsilon_2]$. An individual student's tastes are correlated across majors within a college, and across colleges given the same major. Notice that the consumption value of a major $v_m(x, \epsilon_{1m})$ enters both (1) and, as shown later, one's utility on the job.

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 major-specific efficiency η_m, and chooses to stay or to drop out at the end the first period in college. Student choice is restricted to be between staying and dropping out, which is consistent with the Chilean practice mentioned in the background information section. Later in a counterfactual experiment, we explore the gain from more flexible transfer policies.[17]

---

[15] 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.

[16] Recall that $a \in x$.

[17] We also assume that an enrollee fully observes her efficiency in her major by the end of the first period (2 years) in college. It will be interesting to allow for gradual learning. Given the lack of information on student performance in college, we leave such extensions to future work.

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.

Information Set: Sys.J

Stage	Student	Researcher
1: Enrollment	x, ϵ	x
2: Stay/Drop out	x, ϵ, η_m	x

3.2 Student Problem

This subsection solves the student's problem backwards.[18]

3.2.1 Continuation Decision

After the first college period, an enrollee in (j, m) observes her major-m efficiency η_m, 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.[19] Given peer quality A_{jm}, a student's second-period problem is

$$u_{jm}(x, \epsilon, \eta_m | A_{jm}) = \max \left\{ v_{jm}(x, \epsilon, A_{jm}) + \sum_{\tau'=3}^{T} \beta^{\tau'-2} \left[E_\zeta \left(w_m(\tau - 3, x, \eta_m, A_{jm}, \zeta) \right) + v_m(x, \epsilon) \right], V_d(x) \right\}.$$

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

[18] 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.

[19] Ideally, one would model the dropout and the outside options in further detail, by differentiating various choices within the outside option: 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. 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. See Appendix A1 for functional forms.

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 admitted to and the outside option with value $V_0(x)$, i.e.,

$$U(x, \epsilon|a^*, A) = \max \left\{ \max_{(j,m)|a_m \geq a^*_{jm}} \{\beta E_{\eta_m}(u_{jm}(x, \epsilon, \eta_m|A_{jm})) + v_{jm}(x, \epsilon, A_{jm})\}, V_0(x) \right\}.$$

Let $\delta^1_{jm}(x, \epsilon|a^*, A) = 1$ if program (j, m) is chosen in Stage 1.[20]

3.3 Sorting Equilibrium

Definition 1 *Given cutoffs a^*, a sorting equilibrium consists of a set of student enrollment and continuation strategies $\left\{\delta^1_{jm}(x, \epsilon|a^*, \cdot), \delta^2_{jm}(x, \epsilon, \eta_m|\cdot)\right\}_{jm}$, and the vector of peer quality $A = \{A_{jm}\}_{jm}$, such that*[21]
(a) $\delta^2_{jm}(x, \epsilon, \eta_m|A_{jm})$ is an optimal continuation decision for every (x, ϵ, η_m);
(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),

$$A_{jm} = \frac{\int_x \int_\epsilon \delta^1_{jm}(x, \epsilon|a^*, A) a_m dF_\epsilon(\epsilon) dF_x(x)}{\int_x \int_\epsilon \delta^1_{jm}(x, \epsilon|a^*, A) dF_\epsilon(\epsilon) dF_x(x)}. \tag{2}$$

Finding a sorting equilibrium can be viewed as a classical fixed-point problem of an equilibrium mapping from the support of peer quality A to itself. Online Appendix B5 proves the existence of an equilibrium in a simplified model. Appendix A3 describes our algorithm to search for equilibria, which we always find in practice.[22]

[20] For a student, the enrollment choice is generically unique.

[21] 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. We also need to model how the admissions policies are chosen under Sys.S, which we do in the counterfactual experiments.

[22] Uniqueness of the equilibrium is not guaranteed. Our algorithm deals with this issue using the fact that all equilibrium objects are observed in the data.

3.4 Further Details of the Model

3.4.1 Student Characteristics

There are two family income groups $y \in \{low, high\}$.[23] The major-specific (pre-college) ability is given by

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

where $s = [s_1, s_2, ..., s_S]$ is the vector of test scores that summarize one's knowledge in subjects such as math, language, social science and science; $\omega_m = [\omega_{m1}, ..., \omega_{mS}]$ 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 math knowledge more and language knowledge less than 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.[24] Let M_a be the set of majors within the general interest of a student with ability vector a.[25]

3.4.2 Major-Specific Consumption Values and Education Costs

The per-period consumption value of major m depends on one's major-specific taste and one's characteristics. An individual with higher ability a_m may find it more enjoyable (less costly) to study in major m and work in major-m related jobs. We also allow preferences to differ across genders: on average, some majors may appeal more

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

[24] 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.

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

$$M_a = \{m \in \{1, ..., M\} : a_m \neq n/a\}.$$

We adopt the convention that one's taste $\epsilon_m = -\infty$ if $a_m \notin M_a$.

to females than to males.[26] In particular, $v_m(x, \epsilon_{1m})$ is given by

$$v_m(x, \epsilon_{1m}) = \overline{v}_m I(female) + \lambda_{1m} a_m + \lambda_{2m} a_m^2 + \epsilon_{1m},$$

where the mean major-specific consumption values for males are normalized to zero, and $\overline{v}_m$ is the mean major-m value for females. λ_m's measure how consumption values in major m change with major-specific abilities.[27]

The monetary and effort costs of attending program (j, m), governed by the cost function $C_{jm}(x, A_{jm})$, depend on student characteristics x and peer quality A_{jm}. In particular, we allow the same tuition level to have different cost impacts on students from different family income groups, so as to capture possible credit constraints.[28] The cost also depends on one's own ability $a \in x$, as well as peer quality A_{jm}. For example, it may be more challenging to attend a class with high-ability peers because of direct peer pressure and/or curriculum designs that cater to average student ability.

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. Besides multiple years of macro data, we also obtained micro-level data for the 2011 freshmen cohort. There were 247,360 PSU test takers in 2011. 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] From the 159,365 students, we draw 10,000 students as our final sample due to computational

[26] Gender-specific preferences may arise from not only individual tastes, but also social norms and other channels. We label the combination of all these potential factors as "gender-specific tastes."

[27] In the estimation, we restrict λ_{2m} to be the same across majors.

[28] If we had information on how students financed their college education, we would have modeled credit constraints more explicitly.

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

considerations.[30,31]

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 the worker's 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 annual wage as 12 times the monthly wage, measured in thousands of deflated pesos. For each major, we trimmed 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 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. 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. To obtain information on wages at higher experience levels, we use cross-sectional data from the Chilean Characterization Socioeconomic Survey (CASEN),

[30] For each parameter configuration, we have to solve for the equilibrium via an iterative procedure as discussed in the appendix. Each iteration involves numerically solving the student's problem and integrating out unobserved tastes. This has to be repeated for every student in the sample, since each of them has a different x.

[31] 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).

[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.

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.

Our last data source is the Indices database from the Ministry of Education of Chile. It contains information on college-major-specific tuition, weights used to form the admission score index $(\{\omega_m\})$, the admission cutoffs $\left(\{a^*_{jm}\}\right)$, and the numbers of enrollees in consecutive years.

4.2 Aggregation of Academic Programs

For both sample size (of the wage data) and computational reasons, we have aggregated specific majors into eight categories according to the area of study, coursework, PSU requirements and average wage levels.[33] The eight aggregated majors are: Business, Education, Arts and Social Sciences, Sciences, Engineering, Health, Medicine and Law.[34] We also aggregated individual traditional universities into three tiers based on admissions criteria and student quality.[35] Thus, students have 25 options, including the outside option, in making their enrollment decisions.[36]

Table 1 shows some details about the aggregation. 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 last two columns show similar statistics for total enrollment and tuition. Cross-tier differences are clear: higher-ranked colleges have better students, larger enrollment and higher tuition.

[33] 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.

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

[35] 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.

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

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

4.3 Summary Statistics

This subsection provides summary statistics for the aggregated programs based on our final sample. 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.[37] Compared to average students, females (53% of the sample) are less likely to enroll in college (25% v.s. 28%) and a larger fraction of female enrollees are enrolled in the lowest tier (35% v.s. 32%).

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.[38] This rank is also roughly consistent with the rank 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 extreme. Comparative advantages differ across majors. For example, law and social

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

[38]See Figure 1 in the online appendix for wage paths by major.

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.[39] Θ includes parameters governing the distributions of student tastes, 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

[39] 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 relatively large sample size.

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.[40]

5.1 Target Moments

The combined data sets contain information on various predictions of the model, based on which we choose our target moments. 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.

The PSU data contains information that summarizes the sorting equilibrium: program-specific enrollment and peer quality (Moments $1\,(a)$ and $2\,(a)$ as listed below). It also provides information critical for the identification of student preferences and costs. The different enrollment choices made by students with different demographics (Moments $1\,(a)$) reveal information on the effects of these characteristics on preferences and costs. Students also differ in their unobservables. 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 v.s. 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 $2\,(b)$) is informative about the relationship between peer quality and effort costs. For example, if the relationship is too weak (strong), then more (fewer) students who are eligible will be drawn to programs with better peers in order to benefit from the positive peer effects on wages, which will increase (reduce) the dispersion of abilities in these programs.

In the wage data, the relationship between wages and student's observable charac-

[40]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. The results from the alternative estimation approach are available upon request.

teristics (Moments $4\,(b)$ and $4\,(c)$) provides key information about major-specific human capital production and wage functions. College retention rates (Moments $2\,(c)$), the ability difference between enrollees and those who stayed (Moments $2\,(a)$ and 3), together with the dispersion of wages among workers with similar observables (Moments 4), inform us of the dispersion of major-specific efficiency shocks. For example, lower dispersion in those shocks would lead to higher retention rates and lower wage dispersion; moreover, since pre-college ability is relatively more important in this case, conditional on retention rate, the ability difference between enrollees and graduates should be larger. Finally, Moments 5 inform us of wage growth over the life cycle. In total, we estimate 88 free parameters by matching 448 moments.[41]

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).
    - (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).

[41] 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."

(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 $\tau = 10, ..., 40$.

6 Results

6.1 Parameter Estimates

This section reports the estimates of parameters of major interest. Tables A2.1-A2.4 in the appendix report the estimates of other parameters. Standard errors (in parentheses) are calculated via bootstrapping.[42]

Table 4 Human Capital Production

	Peer Ability (γ_{1m})		Own Ability (γ_{2m})	
Medicine	0.01	(0.002)	0.18	(0.04)
Law	0.58	(0.04)	1.26	(0.02)
Engineering	0.70	(0.01)	1.53	(0.01)
Business	1.48	(0.01)	1.52	(0.01)
Health	0.53	(0.03)	0.48	(0.03)
Science	1.44	(0.01)	1.62	(0.01)
Arts&Social	0.91	(0.02)	1.03	(0.03)
Education	1.08	(0.01)	0.55	(0.02)
Efficiency Shock (σ_η)			0.60	(0.08)

The first eight rows of Table 4 show the estimates of the parameters in the human capital production function, which also measure the elasticities of wages with respect to peer ability and own ability. The left panel shows significant differences in the importance of peer ability across majors: the elasticity of wage with respect to peer

[42] 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 200 bootstrap iterations. Given the sample size (10,000) and the sampling scheme described in footnote 30, the precison of most of our estimates is high.

quality is over 1.4 in business and science, while only 0.01 in medicine.[43] Considering both the left and the right panels of Table 4, we find that the relative importance of peer ability versus own ability differs systematically across majors although no restriction has been imposed in this respect. In 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, while the opposite is true for education, the major with the lowest average wage.[44] This finding has major 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.

The last row of Table 4 shows the dispersion of major-specific efficiency shocks. To understand the magnitude of this estimate (0.6), imagine two counterfactual scenarios: first, if a student's fit to her major were improved by one standard deviation, her starting wage would increase by about 40% ceteris paribus. Second, if the dispersion of these shocks were reduced by 25% from 0.6 to 0.45, the overall college retention rate would increase from 75% in the baseline to about 85%. Clearly, students face non-trivial uncertainties over their major-specific fits.

Table 5 reports parameter estimates for major-specific consumption values. The first two columns show how these values vary with own ability and peer ability. The three majors with highest average wages and social science major are the most satisfying for high ability individuals. Except for engineering, effort costs in these majors are also the most responsive to peer abilities. This is especially true for law programs, which constantly put students in competitive situations such as those in case studies. Empirically, the high cost helps to explain why some law-eligible students chose other majors despite of the expected high wage for law students. Similarly, the consumption value in education is found to be low because the low wages in education are not

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

[44] One possible reason for this finding is statistical discrimination on the labor market. For example, in law and medicine, the practice of licensing and residency/intership reduces the need for statistical discrimination, while the opposite may be true for the education major since the productivity of a potential teacher is difficult to judge from one's own characteristics.

sufficient to explain why most students who were eligible for the education major (the least selective major) chose other majors. The last column of Table 5 shows that on average, females have higher tastes for the conventionally "feminine" majors: health and education, but lower tastes for all the other majors. In the online 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.[45]

Table 5 Consumption Value (Major-Specific Parameters)

	Own Ability		Peer Ability		Female	
Medicine	6.33	(0.73)	-6.11	(0.62)	-1982.2	(255.8)
Law	2.13	(0.41)	-17.46	(1.30)	-196.9	(66.8)
Engineering	2.22	(0.10)	-0.01	(0.002)	-1719.5	(59.7)
Business	0.004	(0.006)	-2.52	(0.20)	-196.6	(33.5)
Health	0.006	(0.004)	-0.24	(0.02)	1668.6	(26.6)
Science	0.001	(0.003)	-0.001	(0.001)	-376.6	(30.6)
Arts&Social	1.50	(0.48)	-4.14	(0.65)	-393.3	(19.4)
Education	0.003	(0.008)	-0.001	(0.02)	1302.5	(17.9)

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.[46] The model slightly underpredicts the fraction of students enrolled in the top tier.

Table 6 Enrollment by Tier (%)

	All		Females	
	Data	Model	Data	Model
Tier 1	5.1	4.5	4.5	3.4
Tier 2	14.1	14.7	12.2	12.1
Tier 3	9.0	9.9	9.1	8.8

[45] 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.

[46] The fits of enrollment patterns for students with low family income are in the online appendix.

Table 7 shows the distribution of enrollees across majors. The fit for the distribution among all enrollees is very close. For female enrollees, the model underpredicts the fraction in social sciences and overpredicts that in education.

Table 7 Enrollee Distribution Across Majors (%)

	All		Females	
	Data	Model	Data	Model
Medicine	3.4	5.0	3.2	2.9
Law	4.6	3.9	4.8	3.6
Engineering	36.6	36.5	23.4	24.2
Business	9.9	9.9	10.5	10.6
Health	11.7	10.7	17.1	17.9
Science	8.5	9.0	8.3	8.0
Arts&Social	11.2	11.0	14.1	10.8
Education	14.0	14.1	18.6	21.8

Table 8 Ability & Retention (by Tier)

	Ability[a]		Retention (%)	
Tier	Data	Model	Data	Model
1	701	701	79.3	79.6
2	624	626	76.5	75.5
3	581	583	68.1	73.2

[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	727	87.6	87.0
Law	658	649	81.3	80.8
Engineering	623	625	71.8	74.4
Business	619	619	74.6	73.4
Health	641	636	79.8	78.0
Science	622	614	63.7	72.0
Arts&Social	612	597	74.3	75.1
Education	590	592	77.1	73.8

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

Table 8 (Table 9) shows the fits of average student ability and retention rates by tier (major).[47] All ability measures are closely matched. The retention rate is over-predicted for Tier 3 in Table 8 and for science in Table 9.

7 Counterfactual Policy Experiments

We first introduce two counterfactual admissions regimes, Sys.S and a hybrid of Sys.S and Sys.J, providing overall cross-system comparisons. Then, we conduct a milder policy change that allows students one chance to switch programs within Sys.J. Finally, we examine the effects of admissions systems in detail, focusing on the contrast between Sys.J (the baseline) and Sys.S.

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, one who aims at maximizing total student welfare by setting college-specific, rather than college-major-specific, admission policies.[48] The constraints for the planner include: 1) a student admitted to a higher-tier college is also admitted to colleges ranked lower, and 2) the planner can use only ability a to distinguish students. These two restrictions keep our counterfactual experiments closer 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.

[47] 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 be retained in (j, m) if she stays in any specific program within our (j, m) category.

[48] The planner takes into account tuition and effort costs for the student in her optimization problem. 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.

Stage 2: Students make enrollment decisions, choosing one of the colleges they are admitted to or the outside option.
Stage 3: An enrollee takes courses in majors within her general interests and learn her efficiency levels in these majors. Then, she chooses one of these majors or drops out.[49]
Stage 4: Stayers spend one more period studying in the major of choice and then enter the labor market.

Information and Decision: Sys.S

Stage 1		Stage 2		Stage 3	
Info	Planner Choice	Info	Student Choice	Info	Student Choice
a	Admissions	x, ϵ	College (j)	$x, \epsilon, \{\eta_m\}_{m \in M_a}$	One major/Dropout

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 expectation of the student's enrollment and major choices, integrating out the student's characteristics and tastes that are unobservable to the planner, and the major-specific efficiency shocks. Then, the planner calculates the expected value for this individual and her effect on peer quality. Peer quality matters both because it affects the market return and because it affects student effort costs. Overall, the planner's optimal admissions policies lead student sorting toward the maximization of total student welfare.[50] Online Appendix B3 contains formal theoretical details.

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.[51] 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, one in this section and another in

[49] In this section, students are free to choose majors. Section 7.2 will explore a case where additional restrictions are imposed.

[50] As a caveat, our policy experiments hold the wage functions unchanged. A more comprehensive model would consider the reactions of labor demand to the new regime, which is beyond the scope of this paper.

[51] 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.

Section 7.2. In each set, we conduct a series of experiments, solving for new equilibria to compare with the baseline.

In this section, we assume that to make up for the first period (2 years) of college spent without specialization, students have to spend, respectively, 0, 1 and 2 extra year(s) in college. Table 10 shows the equilibrium enrollment, retention and student welfare under the baseline and under Sys.S with different lengths of college life. In all cases, postponing major choices increases the overall retention rate from 75% to around 90% : a significant fraction of dropouts occur in the current system because of student-major mismatches.[52] In the first counterfactual case, enrollment increases from 29% to 39%, and the mean student welfare increases by about 4.6 million pesos or 5%. When one has to spend extra time in college, college enrollment decreases sharply. In the last case where the first period of college contributes only to a student knowledge about her major-specific fits but not to her marketable skills, the new system causes a 0.9% welfare loss relative to the baseline. However, we believe the last case to be overly pessimistic: one is likely to obtain at least some basic skills by taking first-year courses even without specialization.

Table 10 Enrollment, Retention & Welfare: Sys.S

	Baseline	0 Extra Year	1 Extra Year	2 Extra Years
Enrollment (%)	29.1	39.1	27.5	19.2
Retention (%)	75.3	91.1	89.2	90.2
Mean Welfare (1,000 Peso)	93,931	98,574	95,185	93,093

7.1.2 Hybrid of Sys.J and Sys.S

Although Sys.S allows students the opportunity to better learn about themselves before choosing their majors, the extra time cost may outweigh the benefit for some students.[53] We therefore consider a hybrid system that combines the merits of Sys.J and Sys.S by allowing students the choice between early and postponed specialization. This hybrid framework involves the following stages:

Stage 1: The planner announces college-specific admissions policies, subject to the same constraints as in Sys.S.

Stage 2: Students make enrollment decisions. An enrollee chooses between claiming

[52] If students face uncertainties other than major-specific efficiency shocks, for example, shocks that change the value of college in general, then we might over-predict the retention rate in the new system.

[53] A discussion about gainers and losers will be provided in the next section.

a major upon enrollment (specializer), or taking courses in majors within her general interests (diversifier).

Stage 3: A specializer learns about her fit in her major and decides whether or not to drop out. A diversifier learns her fits in various majors she has been exposed to and chooses one of these majors or drops out.

Stage 4: Specializers who chose to stay spend one more period specializing before entering the labor market. Diversifiers who chose to stay spend n more years specializing before entering the labor market, where n may be longer than one period.

Information and Decision: Hybrid

Stage 1		Stage 2		Stage 3	
Info	Planner Choice	Info	Student Choice	Info	Student Choice
			Non-Enrollee	x, ϵ	-
a	Admissions	x, ϵ	Specializer (j, m)	x, ϵ, η_m	Stay/Dropout
			Diversifier (j)	$x, \epsilon, \{\eta_m\}_{m \in M_a}$	One major/Dropout

Again, we consider three cases under the hybrid policy, where a diversifier has to spend 0, 1 or 2 extra year(s) in college, compared to a specializer. The results are shown in Table 11. For comparison, we list results from the baseline, the hybrid, as well as the results under Sys.S. Combing the merits of both systems, the hybrid policy leads to greater welfare gains compared to Sys.S: the higher the time cost for diversifiers, the larger the advantages of the hybrid policy. As more students are attracted to colleges under the hybrid policy, retention rates are lower than those under Sys.S in all cases.

Table 11 Enrollment, Retention & Welfare: Sys.S v.s. Hybrid

	Baseline	0 Extra Year		1 Extra Year		2 Extra Years	
		Sys.S	Hybrid	Sys.S	Hybrid	Sys.S	Hybrid
Enrollment (%)	29.1	39.1	40.8	27.5	36.2	19.2	34.7
Retention (%)	75.3	91.1	89.2	89.2	80.1	90.2	75.9
Welfare (1,000 Peso)	93,931	98,574	98,857	95,185	96,098	93,093	95,664

How many students choose to specialize early? As the first row of Table 12 shows, it depends critically on the cost of diversification. When diversification involves no extra time, only 15% enrollees will give up the free opportunity to learn more about themselves. At the other extreme, when two extra years are at stake, 97% of enrollees will choose early specialization. Although it may be costly, diversification does improve a student's chance to find the right match: college retention rates for diversifiers are significantly higher than those for specializers in all cases (Row 2 of Table 12).

Table 12 Enrollment, Retention & Welfare: Hybrid

	0 Extra Year		1 Extra Year		2 Extra Years	
	Specializer	Diversifier	Specializer	Diversifier	Specializer	Diversifier
% of Enrollees	15.2	84.8	69.0	31.0	96.9	3.1
Retention (%)	79.8	90.9	76.3	88.6	75.5	88.8

7.1.3 Rematch Under Sys.J

Although the same rigid transfer policies are practiced in quite some countries like Chile and those described in Appendix B4, 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.[54]

Stage 2: A college enrollee in major m observes her major-specific efficiency η_m, and chooses to stay, to transfer to a different college-major pair, 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 major-specific efficiency in 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 0 and 1 extra year scenarios, rematch policy leads to smaller improvement over the baseline than Sys.S does. When two extra years are required, the rematch policy still improves welfare over the baseline, while Sys.S leads to a welfare loss.

[54] On the student side, we impose the equilibrium peer quality conditions. On the admissions side, we impose the same admissions policies used in the current Chilean system. Results from this experiment are subject to these exogenous admissions policies.

Table 13 Enrollment, Retention & Welfare: Sys.S v.s. Rematch

	Baseline	0 Extra Year		1 Extra Year		2 Extra Years	
		Sys.S	Rematch	Sys.S	Rematch	Sys.S	Rematch
Enrollment (%)	29.1	39.1	32.4	27.5	29.5	19.2	29.4
Retention (%)	75.3	91.1	87.7	89.2	82.1	90.2	78.2
Welfare (1,000 Peso)	93,931	98,574	95,651	95,185	94,418	93,093	94,151

7.2 A Closer Look

We make a more detailed investigation into the impacts of changes in admission policies, focusing on the contrast between the baseline Sys.J and Sys.S.

7.2.1 Gainers and Losers

Who is likely to gain/lose when Sys.J switches to Sys.S? To make a more informative comparison, we hold the average student welfare equalized between the two systems. To do so, instead of extending college life for all, we take an arguably more realistic approach and treat majors differently.[55] For the two most specialized majors, law and medicine, students have to spend more time in college in order to make up for the early non-specialization period. For other majors, the lengths of studies are unchanged at the cost of potential losses of human capital, the production of which becomes $(1-\phi)\, h_m(a_m, A_{jm}, \eta_m)$. Thus, ϕ is the fraction of human capital lost ceteris paribus. Given this framework, we seek the combinations of extra year (in law and medicine) and ϕ (in other majors) under Sys.S that yield the same average student welfare as Sys.J. The results are shown in online Appendix B3.[56] We find that males and students from low income families are more likely to be gainers than their counterparts, and that when a student already has a clear comparative advantage as reflected in her pre-college abilities, the cost of delayed specialization is likely to outweigh its benefit.

[55]For example, in the U.S., for most majors, students receive specialized training only in upper college years. For law and medicine, specialization usually starts after one has received more general college training and lasts another 3 to 6 years.

[56]Each of the following combinations will equalize the welfare: 1) law and medicine majors extend for 1 year, and $\phi = 23\%$ for other majors; or 2) law and medicine majors extend for 2 years, and $\phi = 19.5\%$ for other majors.

7.2.2 Enrollment and Major Choice Distribution

To compare the distributions of student choices, we hold the total enrollment equalized between Sys.S and Sys.J, which happens when medical and law students spend one more year in college and $\phi = 8.5\%$ for other majors. Tables 14-16 are based on this configuration.

Table 14 displays enrollment and retention rates by tier. Compared to the baseline case, Sys.S features more students enrolled in both the top tier (Tier 1) and the bottom tier. What explains the growth of Tier 1 relative to Tier 2? 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 Tier 2 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.[57] 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.

What explains the growth of Tier 3 relative to Tier 2? Although the total enrollment remains the same, the composition of enrollees changes as the system shifts. On the one hand, some former outsiders choose to enroll given the prospect of a better match. A large fraction of them are students with relatively low ability, whom are deemed suitable only for hence admitted only to Tier 3 by the planner. On the other hand, some former enrollees choose the outside option because of the potential loss of either time or human capital ($\phi = 8.5\%$). Since one's outside value increases with one's ability, a lot of students in this group are former Tier 2 enrollees who have middle-level abilities.

Table 14 Enrollment and Retention (%)

	Baseline		New	
	Enrollment	Retention	Enrollment	Retention
Tier 1	4.5	79.6	5.1	93.6
Tier 2	14.7	75.5	12.2	92.5
Tier 3	9.9	73.2	11.7	89.3
All	29.1	75.3	29.1	91.4

[57] Some of these students will opt for a lower-ranked tier due to tastes.

Table 14 also shows that retention rates in all three tiers improve significantly with the change of the system. In fact, even the worst case under the new system (Tier 3) features a retention rate that is 10% higher than the best case under the old system (Tier 1).

Table 15 displays the distribution of students across majors in the first and second period in college.[58] Focusing on the first four columns, we see that without major-specific barriers to enrollment, the fraction of students increases significantly in law and medicine majors. However, enrollment in these two majors are often strictly rationed regardless of the admissions system. We mimic such rationing by adding one more constraint to Sys.S: among all enrollees in college j, only those with law-specific (medicine-specific) ability that meets a certain cutoff have the option to major in law (medicine). We conduct a series of experiments with different cutoffs and report results from the one where the final number of students in each law (medicine) program equals the number of available slots as proxied by the enrollment size of the corresponding program under the baseline.

Table 15 Distribution Across Majors (%)

	Baseline		New		Rationed New	
	1st Period	2nd Period	1st Period	2nd Period	1st Period	2nd Period
Medicine	1.5	1.3	-	3.3	-	1.5
Law	1.1	0.9	-	1.6	-	1.1
Engineering	10.6	7.9	-	7.3	-	7.2
Business	2.9	2.1	-	3.4	-	3.5
Health	3.1	2.4	-	2.8	-	2.6
Science	2.6	1.9	-	3.6	-	3.5
Arts&Social	3.2	2.4	-	2.1	-	2.1
Education	4.1	3.0	-	2.5	-	2.5
All	29.1	21.9	29.1	26.6	26.5	24.1

The last two columns of Table 15 show the equilibrium enrollment with rationing. By construction, the fraction of students majoring in law (medicine) is cut down to its capacity. It is not clear a priori how enrollment in unrationed majors may change

[58] For the first period in college, the distribution across majors is defined only for the baseline case, since in the new system students do not declare majors until the second period.

because two conflicting effects coexist. On the one hand, given total enrollment, enrollments in unrationed majors should increase as rationed-out students reallocate themselves. On the other hand, some students who would enroll without rationing may be discouraged from enrolling at all as they are denied of the option to major in law and medicine. Indeed, as shown in the last row of Table 15, 2.6% fewer students are enrolled in the first period when rationing is imposed. Due to the dominance of this second effect, engineering, health and science majors all become smaller compared to the case without rationing. The only major where the first effect dominates is business, which becomes slightly larger.

7.2.3 Productivity

Table 16 shows the mean log starting wages (in 1,000 pesos) by major, which also reflects the average productivity by major. With or without rationing, Sys.S improves the quality of matches and hence productivity in all majors compared to the baseline. This is true even though there is a 8.5% loss of human capital ceteris paribus for majors other than law and medicine.

Table 16 Log Starting Wage

	Baseline	New	Rationed New
Medicine	9.10	9.17	9.18
Law	9.20	9.59	9.63
Engineering	8.97	9.03	9.03
Business	8.51	8.74	8.76
Health	8.38	8.89	8.90
Science	8.36	9.07	9.08
Arts&Social	8.32	8.79	8.80
Education	8.06	8.35	8.35

When enrollment in law and medicine is rationed, the average productivity increases even further in both majors, which consist of only the very best students. As students who are rationed out of law and medicine reallocate themselves, two conflicting effects occur for the average productivity in other majors. On the one hand, some rationed-out students have higher abilities in multiple majors over an average student, improving the average productivity in the majors they flow into. On the other hand, some rationed-out students are ill suited for other majors, dragging down the average productivity in the majors they flow into. Comparing the last two columns of Table 16, we see that the resulting changes in the productivity of unrationed majors are marginal. However, at least in one major we can see the dominance of the second effect: the major of business gains not only in size (shown in Table 15), but also in average productivity due to the inflow of high-ability students.

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 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, learn about their fits to various majors and then choose their majors. We have shown changes in the distribution of student educational outcomes and provided bounds on potential welfare gains from adopting the new system. We have also explored a hybrid of Sys.S and Sys.J that allows students to choose between early and postponed specialization, which proves to be a very promising admissions policy regime.[59]

Although our empirical application is based on the case of Chile, our framework can be easily adapted to cases in other countries with similar admissions systems. Due to data limitations, we can only provide bounds on the welfare gains from adopting new admissions policies. A natural and interesting extension is to model human capital production explicitly as a cumulative process and to measure achievement at each stage of one's college life. This extension would allow for a more precise estimate of the loss of specific human capital due to delayed specialization and hence a sharper prediction of the impacts on student welfare when the admissions system changes. This

[59] Given the benefit of the hybrid system, one might wonder why a lot of countries choose inflexible systems. This is an important and interesting question that deserves future research.

extension requires information on student performance in college and/or market returns to partial college training. With such data, it is also feasible to relax our assumption about learning speed and model learning as a gradual process where students update their beliefs about their major-specific suitability by observing college performance overtime.[60]

Another extension is to introduce heterogeneity across colleges besides their student quality, which may also affect market returns. One modeling approach is to introduce exogenous college fixed effect, however, as is true for student quality, college "fixed effect" is likely to change with admissions regimes, for example, via instructional investment. Therefore, a more comprehensive model will allow the social planner to choose college investment together with admissions policies. To implement this extension, information on college investment becomes necessary.

Finally, one can also incorporate ex-ante unobserved heterogeneity in student abilities into the framework. The planner will need to infer students' ability from their observed test scores in the making of admissions policies. This extension will be relatively straight forward if the unobserved component of student ability is "private" and does not affect peer quality. In this case, the unobserved ability will play a role similar to student's individual tastes except that the former will directly affect individual's wages. If the unobserved component of ability also contributes to peer quality, then the estimation strategy needs to deal with the fact that the equilibrium objects are no longer observed from the data.

References

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

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.

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

[60] See, for example, Arcidiacono (2004).

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.

Comay, Y.; Melnik, A.; and Pollachek, M. A. (1986) "The Option Value of Education and the Optimal Path for Investment in Human Capital." *International Economic Review* 14(2): 421-35.

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. (2013) "Equilibrium Tuition, Applications, Admissions and Enrollment in the College Market." working paper, University of Wisconsin-Madison.

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. (1989) "Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon." *Economics of Education Review* 8(4): 305-12.

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.

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

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

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

A1 Detailed Functional Form and Distributional Assumptions

A1.1 Cost of college for students:

$$C_{jm}(x, A_{jm}) = p_{jm} + c_1 p_{jm} I(y = low) + c_2 p_{jm}^2 I(y = low) + c_{3m} A_{jm} + c_4 (A_{jm} - a_m)^2,$$

where p_{jm} is the tuition and fee for program (j, m). c_1 and c_2 allow for different tuition impacts on low-family-income student. c_{3m} and c_4 measure the effect of peer quality on effort costs.

A1.2 The value of the outside option and that of dropout depend on one's test scores (s) and one's family income (y). 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:

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

A1.3 Idiosyncratic tastes:

For major: each element in ϵ_1 is independent and $\epsilon_{1m}\tilde{}\,i.i.d.N(0,\sigma^2_{\text{major}})$.

For programs: $\epsilon_{2jm} = \varepsilon_j + \varepsilon_{jm}$, where $\varepsilon_j\tilde{}\,i.i.d.N(\overline{v}_j, \sigma^2_{\text{col}})$ and $\varepsilon_{jm}\tilde{}\,i.i.d.N(0,\sigma^2_{\text{prog}})$. $\overline{v}_j$ is the consumption value of college j for an average student.

A1.4 Log wage function:

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

$\zeta_\tau\tilde{}\,N\left(-0.5\sigma^2_\zeta, \sigma^2_\zeta\right)$ is an i.i.d. transitory wage shock. Elements in the vector η are assumed to be i.i.d. and $\eta_m\tilde{}\,\ln N(-0.5\sigma^2_\eta, \sigma^2_\eta)$.[61]

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

$u_{jm}(x, \epsilon, \eta_m | A_{jm}) =$

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

The adjusted first-period value function is given by

$$\begin{aligned}U(x, \epsilon | a^*, A) &= \max\left\{\max_{(j,m)}\{\beta^2 E_{\eta_m}(u_{jm}(x,\epsilon,\eta_m|A_{jm})) + \sum_{\tau'=1}^{2}\beta^{\tau'-1} v_{jm}(x,\epsilon,A_{jm})\}, V_0\left(x\right)\right\} \\ & \quad s.t.\ E_{\eta_m}(u_{jm}(x,\epsilon,\eta_m|A_{jm})) = -\infty \text{ if } a_m < a^*_{jm}.\end{aligned}$$

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.

[61] Notice that student abilities across majors are correlated, but their efficiency levels are independent across the aggregated majors. We make the independence assumption for identification concerns. Because a student cannot observe η in making enrollment decisions, and because she can only stay in the major of choice or drop out after the realization of η, the correlation between elements in η does not affect her decisions and therefore cannot be identified. If efficiency shocks are positively correlated, we may over-state college retention rates in our counterfactual experiments.

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 $\left\{\delta^1_{jm}\left(x, \epsilon | a^*, A\right)\right\}_{jm}$.[62]

2) Integrate over (x, ϵ) to calculate the aggregate $\{A_{jm}\}_{jm}$ according to (2), thus 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.

[62] Conditional on enrollment in (j, m), 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 $E_{\eta_m}(u_{jm}(x, \epsilon, \eta_m | A_{jm}))$. Details are available upon request.

Additional Tables

1. Data

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.

2. Parameter Estimates

We fix the annual discount rate at 0.9.[63] Table A2.1 shows how the value of one's outside option varies with one's characteristics.[64] The constant term of the outside value for a student from a low income family is only 70% 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 A2.1 Outside Value

Constant (θ_{01})	8919.8	(98.1)
Low Income (θ_{02})	0.70	(0.01)
Language (θ_1)	131.2	(9.8)
Math (θ_2)	133.3	(17.0)
Dropout (ρ)	1.03	(0.01)

Table A2.2 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.[65] The standard deviations of student tastes suggest substantial heterogeneity in student educational preferences.

Table A2.2 Consumption Value (Major-Independent Parameters)

College Value			Major Value		
Tier 1 ($\overline{\upsilon}_1$)	-3311.1	(248.8)	a_m^2 (λ_{2m})	0.011	(0.001)
Tier 2 ($\overline{\upsilon}_2$)	1126.7	(141.1)			
σ_{col}	3197.1	(386.0)	σ_{major}	2344.3	(86.1)
σ_{prog}	1618.5	(242.8)			

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

Table A2.3 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.

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

[64] 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.

[65] 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.

Table A2.3 College Cost (Major-Independent Parameters)

I(Low Inc)*Tuition (c_1)	3.68	(0.19)
I(Low Inc)*Tuition^2 (c_2)	-0.001	(0.0001)
$(a_m - A_{jm})^2$ (c_4)	6.74	(0.61)

Table A2.4 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 across all majors, which contributes to the lower college enrollment rate among females.

Table A2.4 Other Parameters in Log Wage Functions

	Constant		Experience		Experience^2		female	
Medicine	7.78	(0.02)	0.09	(0.003)	-0.002	(0.0001)	-0.37	(0.09)
Law	-2.63	(0.03)	0.11	(0.004)	-0.007	(0.0002)	-0.08	(0.03)
Engineering	-5.38	(0.01)	0.10	(0.001)	-0.002	(0.0003)	-0.19	(0.01)
Business	-10.67	(0.02)	0.11	(0.001)	-0.003	(0.0001)	-0.19	(0.02)
Health	2.30	(0.02)	0.02	(0.002)	-0.0003	(0.0001)	-0.19	(0.02)
Science	-10.94	(0.01)	0.05	(0.001)	-0.0007	(0.0001)	-0.29	(0.03)
Arts&Social	-3.80	(0.01)	0.02	(0.001)	-0.0005	(0.0001)	-0.11	(0.02)
Education	-2.23	(0.02)	0.07	(0.002)	-0.001	(0.0001)	-0.30	(0.04)
Wage Shock (σ_ζ)		0.683		(0.04)				

For Online Publication

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.[66] 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 24.3% to 27.1% (not shown in the Table), it is still lower than that among males (35.9%) . Moreover, compared with males, females are still less likely to enroll in medicine and science and more likely to enroll in social science. One reason is that, on average, males have higher test scores than females; and they have comparative advantage in majors that uses math more than language.[67]

Table B1 Female Enrollee Distribution

(%)	Baseline		New	
	Male	Female	Male	Female
Medicine	6.7	2.9	9.3	6.6
Law	4.0	3.6	3.9	3.6
Engineering	46.3	24.2	45.9	45.7
Business	9.2	10.6	9.2	9.7
Health	4.9	17.9	3.9	4.8
Science	9.8	8.0	9.1	8.4
Arts&Social	11.1	10.8	10.5	12.6
Education	8.0	21.8	8.1	8.5

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

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

B2 Counterfactual Model Details: Sys.S

B2.1 Student Problem

B2.1.1 Continuation Decision

After the first period, a student with ability vector a learn about her fits to majors in her interest set M_a. Given $\left(x, \epsilon, \{\eta_m\}_{m \in M_a}\right)$ and $A_j \equiv \{A_{jm}\}_m$, an enrollee in college j chooses one major of interest or drops out:

$$
\begin{aligned}
& u_j(x, \epsilon, \{\eta_m\}_{m \in M_a} | A_j) = \\
& \max \left\{ \max_{m \in M_a} \left\{ v_{jm}(x, \epsilon, A_{jm}) + E \sum_{\tau'=3}^{T} \beta^{\tau'-2} \left(w_m(\tau - 3, x, \eta_m, A_{jm}, \zeta) + v_m(x, \epsilon) \right) \right\}, V_d\left(x\right) \right\}.
\end{aligned}
$$

Let $\delta^2_{m|j}\left(x, \epsilon, \{\eta_m\}_{m \in M_a} | A_j\right) = 1$ if an enrollee in j with $\left(x, \epsilon, \{\eta_m\}_{m \in M_a}\right)$ chooses major m.

B2.1.2 Enrollment Decision

We assume that in the first period of college, an enrollee pays the averaged cost for and derives the averaged consumption value from majors within her general academic interest.[68] A student chooses the best among colleges she is admitted to and the outside option:

$$
\begin{aligned}
& U\left(x, \epsilon | q\left(a\right), A\right) = \\
& \max \left\{ \max_j \{\beta E_\eta u_j(x, \epsilon, \{\eta_m\}_{m \in M_a} | A_j) + \frac{1}{|M_a|} \sum_{m \in M_a} v_{jm}(x, \epsilon, A_{jm})\}, V_0\left(x\right) \right\} \\
& s.t.\ E_\eta u_j(x, \epsilon, \{\eta_m\}_{m \in M_a} | A_j) = -\infty \text{ if } \psi_j\left(q\left(a\right)\right) = 0,
\end{aligned}
$$

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 admitted to college j. Let $\delta^1_j\left(x, \epsilon | q\left(a\right)\right) = 1$ if the student chooses college j.

B2.2 Planner's Problem

To formalize the constraint on the planner's strategy space, we introduce the following notation. Let $\Xi \equiv \{\chi_1, \chi_2, \chi_3, \chi_4\} = \{[1,1,1], [0,1,1], [0,0,1], [0,0,0]\}$, where the j-th component of each χ_n represents the admissions to college j, i.e., $\chi_{nj} = 1$

[68] Presumably, there will be greater welfare gains if students are allowed more flexibility in their choices of first-period courses. Our results provide a lower benchmark for potential welfare gains from the switch of the admissions system. We leave the extension for future work.

if a student is admitted to college j. Denote the planner's admissions policy for student with ability a as $q(a)$, we restrict the planner's strategy space to be probabilities over Ξ. That is, for all a, $q(a) \in Q \equiv \Delta([1,1,1],[0,1,1],[0,0,1],[0,0,0])$, a convex and compact set. The probability that a student is admitted to college j, denoted as $\psi_j(q(a))$, is given by $\psi_j(q(a)) = \sum_{n=1}^{4} q_n(a)\chi_{nj}$.

Consistent with the assumptions on student course taking, we assume that in the first period in college, a student with interest set M_a will take $\frac{1}{|M_a|}$ slot in each $m \in M_a$, and that in the second period in college, she will take one slot in her chosen major and zero slot in other majors, where $|M_a|$ is the number of majors within the set M_a. Let $z = [y, g]$ be the part of x that is not observable to the planner, the planner's problem reads:

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

where $\widetilde{U}(a|q(a), A) = \int_z \int_\epsilon U(x, \epsilon|q(a), A) dF_\epsilon(\epsilon) dF_z(z|a)$ is the expected utility of student with ability a, integrating out student characteristics that are unobservable to the planner.[69]

For each a, one can take the first order conditions with respect to $\{q_n(a)\}_{n=1}^4$, 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 $q_n(a)$'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:

$$\begin{aligned} & f_a(a) \int_z \int_\epsilon U(x, \epsilon|\cdot) dF_\epsilon(\epsilon) dF_{z|a}(z) \\ & + f_a(a) \sum_j \psi_j(\cdot)\delta_j^1(a|\cdot) \sum_{m\in M_a} \frac{(a_m - A_{jm})}{|M_a|} b_m \gamma_{2m} A_{jm}^{\gamma_{2m}-1} K_{jm} \\ & - f_a(a) \sum_j \psi_j(\cdot)\delta_j^1(a|\cdot) \sum_{m\in M_a} \frac{(a_m - A_{jm})}{|M_a|} \begin{pmatrix} c_{3m}(1 + \sum_{\tau'=1}^2 \beta^{\tau'-1} \frac{\mu_{jm}^2}{\mu_{jm}^1}) \\ +2c_4 \sum_{\tau'=1}^2 \beta^{\tau'-1} \frac{\mu_{jm}^2}{\mu_{jm}^1} (A_{jm} - A'_{jm}) \end{pmatrix}. \end{aligned} \tag{3}$$

[69] 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 n narrowly defined ranges and hence generate n^2 bins of test scores. All a's in the same bin share the same $F_{z|a}(z)$.

Elements in (3) will be defined in the next paragraph. The first line of (3) 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 $\frac{1}{|M_a|}$.

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 each student a takes $\frac{1}{|M_a|}$ seat in major $m \in M_a$. A_{jm} is the average ability among these students.

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

The second line of (3) relates to market return. b_m is the part of expected lifetime income that is common to all graduates from major m.[70] K_{jm} is the average individual contribution to the total market return among students who take courses in (j,m) :

$$\begin{aligned}
K_{jm} &\equiv \frac{\int_a \psi_j(\cdot)I\left(m \in M_a\right)k_{jm}\left(a\right)f_a(a)da}{\mu_{jm}^1}, \\
\text{where } k_{jm}\left(a\right) &= \int_z e^{\alpha_{3m}I(female)}\int_\epsilon \delta_j^1(x,\epsilon|\cdot)\int_\eta \delta_{m|j}^2\left(x,\epsilon,\eta|A_j\right)a_m^{\gamma_{1m}}\eta_m dF_\eta\left(\eta\right)dF_\epsilon(\epsilon)dF_{z|a}\left(z\right).
\end{aligned}$$

Students with higher a_m contribute more to the total market return of their peers. The third line of (3) relates to effort cost. μ_{jm}^2 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

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

$$\begin{aligned}
& b_m \int_z e^{\alpha_{3m}I(female)}\int_\epsilon \delta_j^1(x,\epsilon|\cdot)\int_\eta \delta_{m|j}^2\left(x,\epsilon,\eta|A_j\right)h\left(a_m,A_{jm},\eta\right)dF_\eta\left(\eta\right)dF_\epsilon(\epsilon)dF_{z|a}\left(z\right) \\
= \; & b_m \int_z e^{\alpha_{3m}I(female)}\int_\epsilon \delta_j^1(x,\epsilon|\cdot)\int_\eta \delta_{m|j}^2\left(x,\epsilon,\eta|A_j\right)a_m^{\gamma_{1m}}A_{jm}^{\gamma_{2m}}\eta_m dF_\eta\left(\eta\right)dF_\epsilon(\epsilon)dF_{z|a}\left(z\right) \\
= \; & b_m A_{jm}^{\gamma_{2m}}\int_z e^{\alpha_{3m}I(female)}\int_\epsilon \delta_j^1(x,\epsilon|\cdot)\int_\eta \delta_{m|j}^2\left(x,\epsilon,\eta|A_j\right)a_m^{\gamma_{1m}}\eta_m dF_\eta\left(\eta\right)dF_\epsilon(\epsilon)dF_{z|a}\left(z\right).
\end{aligned}$$

period. Formally,

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

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* $\left\{\delta^1_j(x,\epsilon|q(a),A), \left\{\delta^2_{m|j}(x,\epsilon,\{\eta_m\}_{m\in M_a}|A_j)\right\}_m\right\}_j$, *a set of admissions policies* $\{q^*(a)\}$, *and a set of program-specific vectors* $\{\Omega_{jm}\}_{jm} \equiv \left\{\mu^1_{jm},\mu^2_{jm},K_{jm},A_{jm},A'_{jm}\right\}_{jm}$, *such that*
(a) $\left\{\delta^2_{m|j}(x,\epsilon,\{\eta_m\}_{m\in M_a}|A_j)\right\}_m$ *is an optimal choice of major for every* $(x,\epsilon,\{\eta_m\}_{m\in M_a})$ *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, $\Gamma : O \Rightarrow O$, 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.[71] 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 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.

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

The two combinations that equalize student welfare between the Sys.J and Sys.S are either 1) law and medicine majors extend for 1 year, and $\phi = 23\%$ for other majors; or 2) law and medicine majors extend for 2 years, and $\phi = 19.5\%$ for other majors.

Table B2 Different Treatments Across Majors

	Baseline	Combination 1	Combination 2
Extra Years in Law & Med	-	1	2
ϕ: Loss in Other Majors (%)	-	23.0	19.5
Enrollment (%)	29.1	22.7	22.3
Mean Welfare (1,000 Peso)	93,931	93,934	93,935

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, controlling for their idiosyncratic tastes, drawn from the distribution according to our estimates. Table B3 shows the regression results for Combination 2, the results for Combination 1 are qualitatively similar. Males 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

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

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 B3 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	-0.46**	0.28**	-0.12	1.18*	-0.17	-1.36*	-0.018**
Std. Dev.	0.08	0.07	0.56	0.53	0.44	0.54	0.0035

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

Control for student idiosyncratic tastes.

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

B4. Other Examples of Sys.J[72]

B4.1 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 admitted to a college-major pair if her index is above the program's cutoff.[73]

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).[74]

[72] 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.

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

[74] 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.

B4.2 Japan

1. High School Track: similar to the case in B5.1.

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).

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.

B4.3 Spain

1. High School Track: similar to the case in B5.1, 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.

B4.4 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), religion 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.

B5 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. 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.[75]

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*

[75] 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.

(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* $m \in \{1,2\}$,

$$A_m = \frac{\int_a \int_\epsilon \delta_m(a,\epsilon|A) a_m dF_\epsilon(\epsilon)\, dF_x(x)}{\int_a \int_\epsilon \delta_m(a,\epsilon|A) dF_\epsilon(\epsilon)\, dF_x(x)}. \tag{4}$$

Proposition 1 *A sorting equilibrium exists.*

Proof. The model can be viewed as a mapping

$$\Gamma : O \Rightarrow O,$$

where $O = \left[0, \overline{A}\right]^2$, $o = [A_1, A_2] \in O$.
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 $\epsilon^*(a,A) \equiv v_2(a,A_2) - \max\{0, v_1(a,A_1)\}$

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

Given that both $v_a(a,A_a) \in R$ 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 (4), 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 B4 Enrollment (Low Income) (%)

	Data	Model
Tier 1	2.3	2.6
Tier 2	12.6	12.4
Tier 3	9.7	9.7

Enrollment among students with low family income.

Table B5 Enrollee Distribution Across Majors (Low Income) (%)

	Data	Model
Medicine	1.7	3.0
Law	3.4	3.2
Engineering	35.1	34.8
Business	10.0	9.9
Health	12.2	10.4
Science	8.2	9.6
Arts&Social	11.0	12.6
Education	18.5	16.4

Distribution across majors among enrollees with low family income.

Table B6 Mean Test Scores Among Outsiders

	Data	Model
Math	533	531
Language	532	532
HS GPA	542	541
Max(Science, Soc Science)	531	530

Mean test scores among students who chose the outside option.

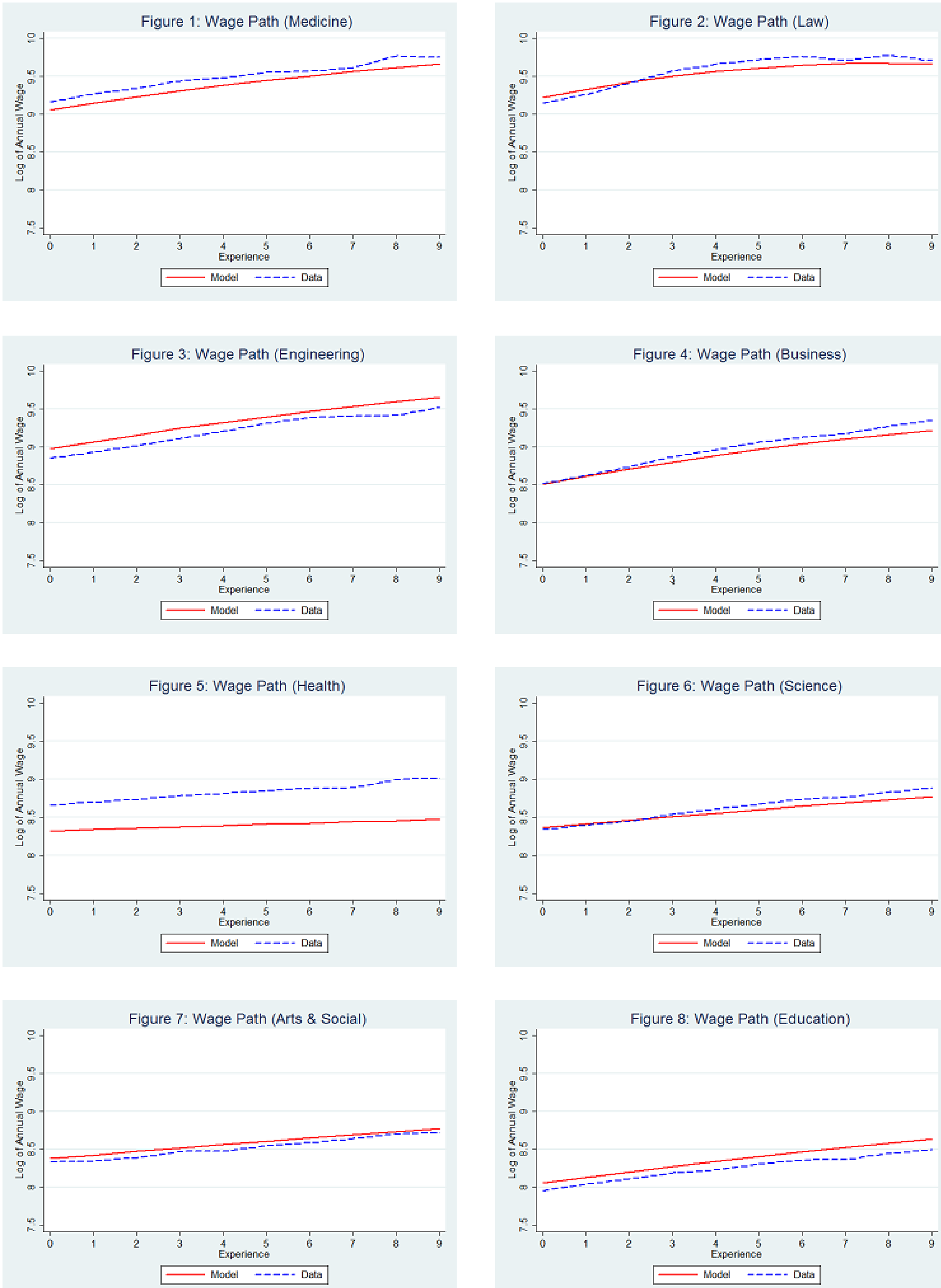

Figure 1: Average Wage by Major and Experience