Quantifying individual uncertainty in decision-making: Unrelated preferences for degree programs reduce students' first-year retention in higher education.

Cristian Candia [1,2], J. Davyt-Colo [3], M. Guevara [4], J. Pulgar [5], T. Yaikin [3], C. Monge [6], F. Pinheiro [7], C. Rodriguez-Sickert [3]

[1]Data Science Institute, Facultad de Ingeniería, Universidad del Desarrollo, Las Condes, 7610658, Chile.

[2]Northwestern Institute on Complex Systems (NICO), Northwestern University, Evanston, IL 60208, USA

[3]Centro de Investigación en Complejidad Social, Facultad de Gobierno, Universidad del Desarrollo, Chile.

[4]Department of Computer Science, Universidad de Playa Ancha, Chile.

[5]Departamento de Física, Universidad del Bío Bío, Concepción 4051381, Chile

[6]Feedback comunicaciones, Vitacura, Chile.

[7]Nova Information Management School (NOVA IMS), Universidade Nova de Lisboa, Lisboa, Portugal

Uncertainty describes the quality of our information concerning risk for a given decision. For those decisions affecting our lives in the long-term, we usually invest time exploring and sampling information to reduce such uncertainty. Higher education is a milestone in people's lives. Therefore, sampling information regarding degree programs is pivotal to reap long-term educational outcomes. We propose a new framework that quantifies individual uncertainty based on a network structure of degree programs, the Higher Education Space (HES). We build the HES using data on 1.6 million applicants' preferences in Chile between 2005 and 2019 (top panel, Figure 1). Then, we quantify individual uncertainty by computing the relatedness of degree program preferences (Rk, Eq. 1) using:

$$R_k = \frac{1}{N(N-1)} \sum_{i \neq j \in P_k} d_{ij}, \qquad (1)$$

where N is the number of degree programs in the application (P_k) of individual k, and d_{ij} is the network distance between degree programs i and j in application P_k. To evaluate the impact of uncertainty in individual decision-making, we test whether the relatedness between applied degree programs impacts first-year retention (the continued enrollment in the same degree program). Note that not-retention is arguably a costly outcome for both individuals and institutions. We find that, on average, students who select related programs have a 73% first-year retention probability, which becomes 40% when applying to unrelated degree programs (bottom panel, Fig. 1). This effect is steeper for high-score applicants (yellow curve, bottom panel, Fig. 1). Indeed, the retention rate equalizes between low and high-score applicants when applying to distant programs. The results are robust to all the available socio-economic control variables such as family income, enrolled institution, province of origin, province of enrolment, and family size, among many others. Besides, we replicate our results using data from Portugal and a regression discontinuity design quasi-experiment for causal inference. Finally, we build prediction models for classifying students in a potential academic risk in an early stage (the beginning of their academic year) with an 80% accuracy. Thus, we provide a network-science framework for quantifying uncertainty in decision-making processes that impact long-term outcomes. In this case, our framework can help prevent academic dropouts by identifying students at risk early and then to focalize institutional accompanying programs such as tutoring, vocational interventions, or remedial classes.

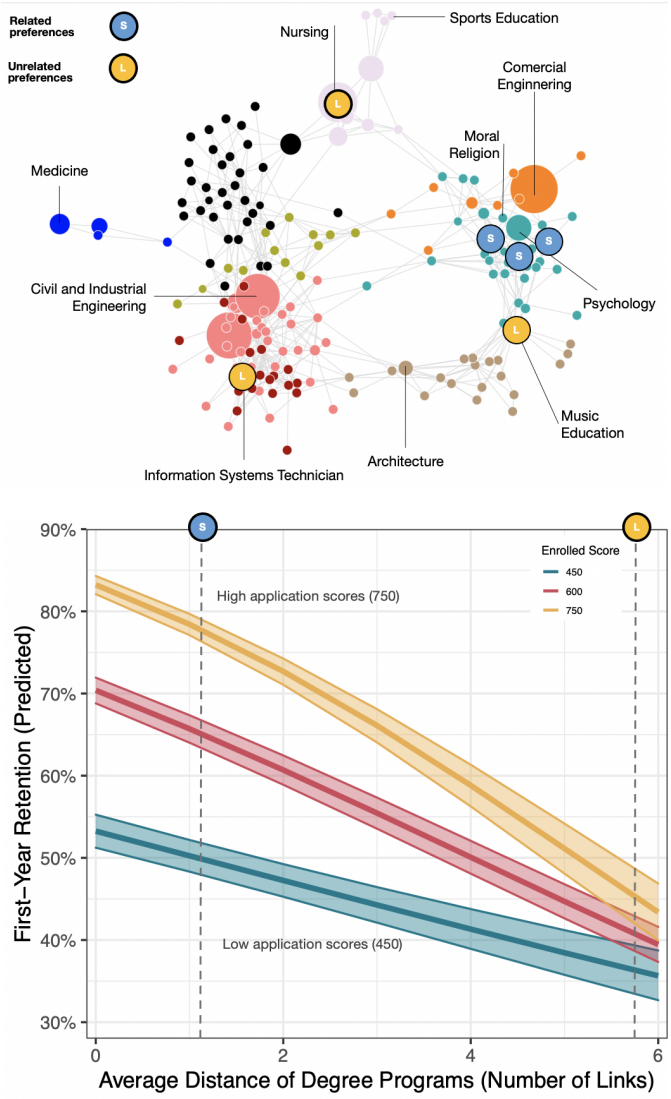

Fig. 1: The top panel shows Higher Education Space (HES) and degree program relatedness. The bottom panel exposes the impact of average distance degree programs on predicted retention probability.