CICS UDD
CENTRO DE INVESTIGACIÓN EN COMPLEJIDAD SOCIAL

The Cost of Being "Late": Tracing Misalignment and Misdiagnosis Through Relational Architectures

Joselina Beatriz Davyt Colo

Doctor en Ciencias de la Complejidad Social
in the Universidad del Desarrollo

Supervisors:
Cristian Candia-Vallejos
Patricia Soto-Icaza

Universidad del Desarrollo
Facultad de Gobierno
Centro Investigación en Complejidad Social

April 20, 2026

The Cost of Being "Late": Tracing Misalignment and Misdiagnosis Through Relational Architectures

Dissertation submitted by **Joselina Beatriz Davyt Colo** (student no. 201908060005S005) in partial fulfillment of the requirements for the Degree of Doctorado en Ciencias de la Complejidad Social in the Universidad del Desarrollo.

Supervisors:

Cristian Candia-Vallejos
Patricia Soto-Icaza

CICS UDD
CENTRO DE INVESTIGACIÓN EN COMPLEJIDAD SOCIAL

Av. Plaza 680, Las Condes. Santiago de Chile, Chile.

April 20, 2026

"La vie est un risque inconsidéré que nous, les vivants, courons. La douceur est d'accepter ce risque, non de fuir l'incertitude"
— Anne Dufourmantelle, Eloge du risque

Agradecimientos

Dedicado a mis padres Norma y Julio, y a mi hija Eloise.

Quiero comenzar agradeciendo a mi familia, pues sin su apoyo constante, emprender este camino hubiera sido imposible. A mis padres, Norma y Julio, por esa lección vital de avanzar siempre y confiar, ante todo, en uno mismo. A mis hermanos, gracias por estar siempre presentes sosteniendo este proceso y por darme el impulso necesario, literal y metafóricamente, para subirme a ese avión y perseguir esta meta. A mis suegros, gracias por cuidar los detalles cotidianos, por esos almuerzos y cenas que me regalaron tiempo y tranquilidad para concentrarme.

A ti, Diego Jauré, mi compañero de vida: gracias por todo. Por el apoyo emocional, por la revisión técnica y por estar incondicionalmente a mi lado en cada paso de esta tesis.

A mi pequeña Eloise, gracias por tu paciencia infinita al compartir a tu mamá con la academia. Gracias por entender las largas jornadas y esperarme; este logro también es tuyo.

Un agradecimiento especial a mis amigos Jaime Salazar, Isabel Mago, Daniela Ortega y Patricio Carvajal. Ustedes fueron mi hogar en tiempos de tormenta. Gracias por abrirme las puertas de su casa durante el estallido social y la pandemia; hicieron que mi estadía en Chile fuera no solo posible, sino cálida y amable.

A mis colegas del CICS, CRISLAB y NEUROCICS, gracias por ser brújula y desafío. Las discusiones compartidas y el enriquecimiento mutuo fueron fundamentales para dar forma a estas ideas.

Por último, mi inmensa gratitud a mis tutores, Cristian y Patricia. Gracias por su profesionalismo, pero sobre todo por su calidad humana. Gracias por la paciencia, la humildad y la disposición para guiarme. Sin su constancia y su excelencia, esta tesis no sería lo que es hoy.

Abstract

The main focus of this thesis is the investigation of the structural origins of temporal disadvantage, manifested as reduced academic persistence and diagnostic delay, across two distinct domains: higher education and clinical diagnosis. Adopting a Computational Social Science framework, we employ network analysis not merely as a tool, but as a theoretical lens to visualize and quantify the relational architectures that generate structural misalignment between individuals and institutions. The first study examines informational inequality in higher education. Using administrative data from 1.6 million applicants in Chile and replicating findings in Portugal, we construct a network of degree preferences to quantify Preference Misalignment—the distance between a student's true interests and their enrolled program. We demonstrate that this topological distance is a robust predictor of first-year retention: students with misaligned preferences face a significantly higher risk of persistency, a penalty that remains even among high-performing students. The second study addresses recognitional inequality in autism. Analyzing a clinical sample of autistic children without intellectual disability, we integrate bipartite and multilayer networks to decode the mechanisms behind the diagnostic delay in autistic girls (averaging two years later than autistic boys). We identify while structural connectivity (i.e., high betweenness and participation coefficient) accelerates diagnosis in boys, it significantly delays it in girls. Conversely, phenotypic entropy (systemic disorder) facilitates recognition in females. Multilayer analysis further revealed that autistic girls exhibit a differentiated physiological and cognitive architecture, in which high cognitive ability is structurally coupled with increased psychological burden. Together, these findings reveal that being "late" is rarely an individual accident, but a consequence of institutional architectures that render certain profiles illegible. By mapping these invisible structures, this work provides empirical evidence for policy interventions aimed at reducing the structural friction that generates inequality.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Being "late" is rarely a neutral condition. In social and clinical trajectories, delay often translates into disadvantage. This disadvantage is seen in the higher education student who realizes too late that her enrollment program misaligns with her preferences. It is also seen in the child whose autistic profile diverges from prototypical diagnostic patterns and is recognized, on average, years later than their peers. The impact of such disadvantages can be profound, potentially altering an individual's life course and development.

To situate, understand, and interpret this phenomenon, its emergence, as well as its potential social impacts, and to serve as a general framework applicable across diverse problem domains, we introduce the concept of:

Definition 1.1 *The temporal disadvantage is comprehensive as the phenomenon of latency in the initialization of a critical trajectory. Conceptually, this refers both to a chronological 'delay', and also to the absence of an 'effective onset' at a moment when the institutional system expects a decision or intervention. This disadvantage arises when structural barriers generate a period of 'drift' or stagnation. During this latency period, the individual does not accumulate expected benefits (such as academic credits or therapeutic interventions) but rather experiences an accumulation of unmet needs that constitutes developmental debt, heightening future vulnerabilities.*

Temporal Disadvantage must be understood as an emergent property of the coupling between institutional architectures, individual profiles, and temporal windows of decision or recognition. Arises as the lack of informational capital in higher-education decisions or the clinical recognition challenge of autism phenotypes. Consequently, temporal disadvantage could acts as a stratification mechanism: postponing the moment of onset alters the developmental trajectory, constituting a gap with the potential to generate structural or permanent inequality.

From this perspective, diagnostic and academic delays are not individual accidents but manifestations of broader structural patterns. Structural patterns are defined as institutional and social processes (the "social layer"), and their consequences are experienced by individuals as cognitive friction, confusion, or a failure of recognition (the "psychological

layer"). These mechanisms generate a "temporal disadvantage", which we understood as a cumulative lag relative to social and developmental milestones.

The temporal disadvantage manifests as lost years of quality of life and identity formation in the absence of a timely diagnosis, or as the friction of being "left behind" peers due to rigid academic pathways that complicate course correction. This loss of time restricts future capabilities and opportunities (Sen, 1992), effectively functioning as a mechanism of stratification (Merton, 1968). Foundational work on inequality has established that such mechanisms, ranging from informational asymmetries (Stiglitz, 2002) to failures of recognition (Fraser, 2000; Honneth, 1995), solidify into forms of "durable inequality" (Tilly, 1998).

Structural patterns are embedded within the very relational architectures that shape social life. A relational architecture refers to the underlying rules of connection that structure a given system. Drawing on recent advances in complex systems, we understand these architectures not merely as sets of pairwise links, but as higher-order structures that capture complex group dependencies and "structural shadows" often missed by traditional models (Battiston et al., 2020). These underlying rules define the actual pathways for flow and function within the system (Lambiotte et al., 2019), actively encoding categorical biases (Noble, 2018; Pendse et al., 2022; Spiel et al., 2022). Such biased architectures impair collective search and problem-solving. By prematurely constraining the exploration of diverse solutions, they affects individuals whose needs or profiles diverge from a predefined norm (Smaldino, 2024).

The friction generated by these biased architectures manifests as structural misalignment. A structural misalignment occurs when the internal rules of an institutional architecture fail systematically to align with the needs, preferences, or profiles of the individuals it is meant to serve. This constitutes a systematic failure where the rules embedded in relational architectures diverge from the needs or profiles of people, crystallizing into a cumulative temporary disadvantage.

The thesis conceptualizes the "cost of being late" (the temporal disadvantage) as a consequence of such structural misalignments. In the educational domain, we trace a misalignment quantified as distance within a network of academic choices, creating a context where students face a greater risk of attrition. In the clinical domain, we identify a divergence from prototypical diagnostic patterns that have traditionally shaped autism recognition and assessment (Gesi et al., 2021; Loomes et al., 2017b),where the child is recognized, on average, years later than their peers. Such misalignments result in the architecture of institutional illegibility, where costly temporal detours are created for those who, by virtue of their sex or background, present a less legible pattern, while streamlining trajectories for those who fit the archetype.

In sum, the thesis focus in two domains characterized by structural misalignment:

- In higher education retention, misalignment is driven by informational friction. Students often enroll with limited knowledge of program requirements or pathways, a systemic issue rooted in an opaque or poorly designed choice architecture that burdens those who do not fit the expected profile.

- In sex autism diferences, the misalignment arises from differential recognition processes. Diagnostic frameworks historically developed around prototypical presentations are structurally less attuned to non-prototypical profiles, contributing to delayed or missed recognition, particularly in females.

A growing body of research has established that outcomes such as student persistence in higher education and delays in clinical diagnosis are shaped by a complex combination of individual characteristics, institutional rules, and contextual constraints (Mandell et al., 2005; Shattuck et al., 2009; Tinto, 1993). This thesis builds directly on that foundational literature by introducing a complementary analytical perspective: relational architectures as an additional layer of information regarding how inequality is produced and sustained.

Rather than focusing exclusively on individual attributes or institutional variables in isolation, this work examines how these elements are organized relative to one another within structured systems. According to the critique often leveled at variable-centered sociology (Emirbayer, 1997). Network-based representations make it possible to formalize this organization, capturing patterns of proximity, clustering, and accessibility that remain implicit in traditional approaches (Lazer et al., 2009; Watts, 2004). Crucially, these relational patterns do not replace existing explanations; they restructure them by revealing how known factors interact within a shared architecture.

The thesis adopts Computational Social Siences (CSS) to give empirical and measurable scope to the theoretical understanding of structural inequality. CSS provides the lens and toolkit to move from variables to relations. Within this framework, network methods in particular make a critical difference. Networks are not just a technical tool but an epistemological bridge (Granovetter, 1973; Watts, 2004) that allows the discovery of hidden patterns and the unveiling of relational architectures. Therefore, the methodological approach of this dissertation is to build one-to-one (pairwise) networks, not as a literal map of reality, but as the measurable "shadow" or projection of that deeper, higher-order architecture. The structures we find in these simpler graphs—their clusters, paths, and communities—are thus interpreted as tangible evidence of those underlying group mechanisms. This allows us to quantify the "topological landscape" of possibilities an individual must navigate, using a tractable method while retaining theoretical depth. The relational turn allows for rethinking inequality as an emergent property of these complex systems.

In the context of higher education (Chapter 1), previous research has extensively documented the role of academic preparation, socioeconomic background, and institutional characteristics in shaping student outcomes (Bourdieu, 1977; Tinto, 1993). The specific contribution of this dissertation is to demonstrate that these well-known factors are embedded in a structured space of degree programs, where the relative position of choices matters. By modeling preferences as a network and measuring their topological coherence, the analysis adds a layer of information about how students navigate constrained choice architectures (Thaler & Sunstein, 2008), offering insights that go beyond what can be inferred from individual test scores or background variables alone.

Similarly, in autism research (Chapter 2), extensive evidence exists regarding sex differences in symptom expression, camouflaging, and diagnostic practices (Hull et al., 2019; M. C. Lai et al., 2017; M.-C. Lai & Szatmari, 2020; M.-C. Lai et al., 2011; Loomes et al.,

2017a). This thesis contributes by showing that these differences are not solely a matter of feature presence or intensity, but also of how clinical characteristics are relationally organized. Network representations of features co-occurrence and integration across domains (Borsboom, 2017) make visible the structural patterns that affect how phenotypic profiles are recognized or missed within diagnostic systems (Cantor et al., 1980; Rosch, 1973).

From this perspective, networks are not treated here as causal mechanisms per se, but as formal descriptions of the structural context in which individual trajectories unfold. Relational architectures define the landscape of similarity, prototypicality, and accessibility that institutions implicitly rely on when making decisions. They shape which profiles are easily recognized, which choices appear coherent, and which trajectories are more likely to be supported or delayed.

The work traces the relational configurations that make delay socially meaningful. In higher education, academic retention challenge exemplifies inequality as informational asymmetry. The work moves beyond prior research by conceptualizing this mismatch as a relational property. A network of university degree programs is generated based on patterns of student application preferences. Misalignment is then operationalized as the preference dissimilarity (network distance) between a student's preference profile and their enrolled program. The findings identify this network-based misalignment as a robust behavioral signal of informational inequality, showing that it imposes an early temporal penalty, significantly decreasing the probability of retention and possible a causal property.

In autism, delays in diagnosis illustrate inequality and constrained recognition. Diagnostic frameworks were historically normed on male phenotypes (Gesi et al., 2021; Loomes et al., 2017b), rendering non-canonical presentations less visible, particularly in females. The study provides evidence of this recognitional inequality. Girls are diagnosed on average approximately two years later than boys. Furthermore, network topology reveals that the cost of illegibility affects both groups, though in distinct ways: possessing a diffuse, non-canonical behavioral profile increases diagnostic delay in boys, eclipsing them for deviating from the standard male prototype. While, surprisingly decreasing it in girls, likely because dispersion in females triggers other external attention. A multilayer network analysis—integrating behavioral data (ADOS-2 items) and cognitive performance (IQ descriptors) with co-occurrence conditions—reveals distinct structural patterns of autistic presentation between boys and girls. These findings suggest that the gendered nature of diagnosis is reflected in divergent topological organizations—in which female networks show higher integration of co-occurring conditions while male networks exhibit greater modularity—rather than just differences in counts of autism-related behavior.

Taken together, these results suggest that retention challenges and diagnostic delay, while seemingly distinct, can be theorized within a shared framework of inequality: temporal disadvantage, that arises from phenomena like informational asymmetry, constrained recognition, and restricted access. By situating our empirical analyses within these theoretical traditions, and by employing computational network methods as a common language, we highlight the relational architectures through which inequality is produced, sustained, and revealed across domains.

By adding the relational layer, the thesis extends the explanatory scope of established

findings. It provides a common language to describe how informational inequality in higher education and recognition constrains in clinical diagnosis emerge from the interaction between individuals and the architectures of the systems they encounter. In doing so, it highlights how temporal disadvantage arises not solely from individual deficits or institutional failures, but from their alignment or misalignment within structured relational spaces.

The empirical evidence presented here carries significant implications for public policy and design. Interventions aimed at reducing educational inequality should take a broader approach. Crucially, recent evidence demonstrates that large proportions of students make "application mistakes," and that interventions focusing on re-architecting the choice system (e.g., changes to information policy design) significantly reduce these mistakes and improve matching outcomes. However, recognizing that institutional architectures often possess a high degree of inertia, the immediate challenge lies in enhancing the legibility of these systems. Consequently, the focus should be on designing robust informational interfaces that help individuals navigate existing structures, ensuring that guidance mechanisms are sufficiently aligned with the architecture to prevent costly mismatches.

In the clinical domain, the proposed network approach offers a complement to existing tools, helping clinicians become more attuned to structural constrained and the particular configurations of non-prototypical profiles, especially in girls. The resulting architecture of illegibility creates a systemic reliance on passive diagnostic pathways, which are historically calibrated toward overt behavioral disruptions. This focus often leaves more socially nuanced or 'adjusted' presentations without the benefit of active, gender-sensitive identification strategies (Loomes et al., 2017b).

Similarly, in the educational sphere, this architecture of illegibility sustains a passive reliance on the student's pre-existing social capital. This is evidenced by the fact that even high-achieving students from disadvantaged backgrounds often fail to navigate selective academic pathways when the choice architecture remains opaque (Hoxby & Avery, 2012). In both domains, the transition from passive to active institutional frameworks is essential to mitigate the structural misalignments that currently penalize those who deviate from the expected archetype.

The thesis is organized as follows: Chapter 2 presents the first empirical study on academic retention and misalignment preferences in higher education. Chapter 3 presents the second empirical study on recognition and diagnostic sex differences in autism. Finally, Chapter 4 provides a general conclusion, synthesizing the findings from both domains, discussing the methodological contributions, and exploring the broader implications for theory and policy.

Chapter 2

Preference Misalignment Reduces Retention in Higher Education

2.1 Abstract

Decision-making under uncertainty plays a central role in shaping individual outcomes, including educational attainment. While prior research has established academic performance, socioeconomic background, and institutional characteristics as key predictors of university retention, the role of students' application strategies–specifically the structure of their stated preferences–remains poorly understood. Here, we develop a network-based measure to quantify individual preference misalignment as an intrinsic attribute of the student's decision-making process. Using administrative data from 1.6 million Chilean university applicants (2012–2021), we show that students with more aligned preferences–those applying to closer programs in the degree programs network–are significantly more likely to persist in higher education. First-year retention reaches 73% among those with the lowest preference distance, compared to 40% for those with the highest. These differences hold across enrollment scores levels, and are most pronounced among high-scoring students. To estimate causal effects, we implement a fuzzy regression discontinuity design exploiting admissions thresholds for first-choice programs. We find that access to the top-ranked program increases retention only when preference alignment is high, revealing an interaction between institutional assignment and individual-level coherence in educational choices. Access to the first-choice program significantly attenuates the risk associated with preference misalignment. Replication using data from Portuguese university applicants confirms the robustness and generalizability of our findings across distinct systems. These results position preference alignment as a key, previously overlooked predictor of student persistence. They highlight the importance of structured guidance in the application process and suggest that improving the coherence of educational choices may enhance retention, reduce attrition, and optimize the impact of public investment in higher education.

Keywords: Higher Education Preferences, Student Retention, Preference Alignment, Incomplete Information, Computational Social Science, Network Science, Quasi-experimental,

Regression Discontinuity.

2.2 Introduction

Decision-making under uncertainty is a central feature of human behaviour, influencing outcomes across domains, including education (Alchian, 1950; De Berker et al., 2016; Kelly, 2001; Kochenderfer, 2015; Pucciarelli & Kaplan, 2016; Tversky & Kahneman, 1974). Particularly in higher education, students face complex decisions when selecting academic programs, often navigating incomplete information, constrained options, and uncertain prospects (Barnett, 2007; Flum & Blustein, 2000; Hastings et al., 2018; Larroucau et al., 2024; Manski, 1991; Manski, 2004; Perez et al., 2014a). These decisions carry high stakes: globally, approximately 31% of tertiary education students fail to graduate, resulting in significant societal and economic costs (Aina et al., 2018). In 2009, dropout rates reached dramatic levels in countries like the U.S. and Italy, where over half of university students abandoned their programs without achieving a degree. In contrast, dropout rates were below 24% in countries such as Belgium, Denmark, and Japan, underscoring significant disparities across regions (for Economic Co-operation & Development, 2009). In Chile, where 28% of students do not complete their degrees, the loss amounts to 30% of public education spending (Gallegos et al., 2018; González & Uribe, 2002).

The literature on student persistence identifies a range of interconnected factors, including academic performance (O'Neill et al., 2011; Stewart et al., 2015; Voelkle & Sander, 2008), family support (Begerson, 2009; Mitchall & Jeager, 2018), and psychological well-being (Brannan et al., 2013; Hamdan-Mansour & Dawani, 2008; Romero et al., 2015). Socioeconomic status plays a critical role as well, as students from low-income families often face a lack of cultural and social capital needed to navigate the application process (Bourdieu, 1977; Mitchall & Jeager, 2018). For instance, fewer than half of low-income adolescents report receiving assistance with college applications (Cochran & Coles, 2012). This lack of guidance can result in "undermatching," where students select institutions or programs below their academic potential, reducing their chances of success (Belasco & Trivette, 2015; J. Smith et al., 2013). At a broader level, Bourdieu's concept of 'habitus' highlights how students' ingrained dispositions, shaped by their social environments, may not align with institutional expectations, contributing to lower rates of academic persistence (Bourdieu, 1977).

Students in centralized higher education systems (such as in Chile, China, Hungary, or Portugal) are required to rank their preferences on academic programs under conditions of constrained information and choice, influenced by cutoff scores, program characteristics, and career expectations (Santelices et al., 2020). Informational interventions, such as those studied in Chile (Hastings et al., 2018) and China (Ye, 2024), have shown that improving access to program information enhances student-program alignment and retention. The vocational exploration process, encompassing self-reflection and external evaluation of educational opportunities, is critical for developing coherent preferences that align students' internal attributes with program realities (Flum & Blustein, 2000). However, limited access to relevant and timely information often disrupts this process, leading

to poorly informed decisions. This issue is especially pronounced in contexts where students have unequal access to vocational guidance, since structured vocational activities in high school have been shown to enhance student knowledge about higher education, particularly among students of higher income, improving their ability to make informed choices and enroll in programs better aligned with their academic and career goals (Santelices et al., 2025). Mismatches between students' preferences and institutional offerings disrupt academic and social integration, both of which are critical for persistence in higher education (Avery & Kane, 2004; Tinto, 1993).

Preference alignment, measured as the relatedness of students' program choices, emerges as a behavioural signal of how individuals navigate constrained choice spaces, such as higher education systems. Relatedness captures the degree of similarity among chosen programs, reflecting the extent to which preferences align with students' knowledge, aspirations, competencies, and available options. Misaligned preferences may indicate informational gaps, uncertainty, or external pressures, which could hinder decision-making and reduce the likelihood of successful institutional retention. At a systemic level, analyzing patterns of preferences and their degree of coherence provides insights into broader inequities in access to information and opportunities, while also revealing how educational institutions at secondary and tertiary levels shape individual persistence outcomes.

This study examines whether the relatedness of students' program preferences impacts first-year retention. Using data from 1.6 million applicants in Chile (2005–2019), we quantify relatedness through a network-based analysis of degree programs. This method captures systemic trends in preference alignment while showcasing disparities in decision-making processes. A higher relatedness in preferences, as proposed, is the consequence of alignment between students' aspirations and program realities, which can foster academic integration, satisfaction, and reduce the likelihood of attrition. Additionally, we validate these findings in Portugal, a context with a similarly centralized admissions system but distinct cultural and institutional features, to test their generalizability across diverse educational settings.

By framing relatedness as a behavioural proxy for program coherence, this study offers a novel perspective on educational decision-making. Our findings have implications for designing interventions such as tailored career counseling, exploratory first-year programs, and informational campaigns that enhance alignment between students' aspirations and institutional offerings, ultimately improving retention and optimizing public investment in higher education.

2.3 Results

Higher education applicants navigate a complex decision-making process that involves selecting degree programs under uncertainty. To examine the structure of applicants' preferences, we constructed a network of degree programs based on application data from 2012 to 2019, comprising 156 nodes (Fig. 2.1A) that correspond to generic degree programs used in the national higher education system. This network connects pairs of degree programs according to their co-occurrence in applicants' preference sets, allowing us to in-

fer their similarity. We identify pairs of degree programs with positive and statistically significant co-occurrences, quantified using ϕ-correlation (Candia et al., 2019). Negative correlations are discarded, and statistical tests ensure that links are not due to chance, filtering out correlations that are indistinguishable from zero (p-value ≤ 0.05). To maintain analytical clarity, the network is binarized, with links included or excluded based on statistical significance (for more details see methods section 2.4).

The structure of this network reveals strong autocorrelations among key attributes of degree programs, including gender composition, application scores, demand-supply ratios, unemployment rates, retention rates, and geographical mobility (Fig. 2.1, panel B). These correlations indicate that degree programs that are closer in the network tend to share similar characteristics, validating the use of network distance as a meaningful measure of preference alignment. Additionally, a survival analysis of network links over time reveals that a significant proportion of connections between degree programs remain stable across years indicating that the higher education space exhibits stable persistent structural properties rather than short-term fluctuations in student preferences.

Applicants are allowed to list up to 10 degree programs in their applications, yet empirical patterns indicate that most students apply to significantly fewer programs, with a strong concentration in the first five choices. Sensitivity analyses using 3, 5, and 10 preferences confirm that students' preferences exhibit diminishing strength beyond the first five choices, suggesting that after this threshold, preferences may be weakly held or less informative (for more details see methods section 2.4).

Beyond its structural properties, the higher education network allows for the computation of a topological distance measure, which quantifies how closely related a student's preferred degree programs are. This preference distance is defined as the average network distance among the degree programs listed in an individual's application:

$$C_k = \frac{1}{N_k} \sum_{i,j \in P_k} d_{ij}$$

where C_k represents the preference distance of student k, N_k is the number of degree programs in the application P_k, and d_{ij} denotes the network distance between degree programs i and j.

This measure distinguishes between students with highly aligned degree preferences (Fig. 2.1A), who choose closely related programs, and those whose preferences span distant and unrelated fields (Fig. 2.1A). The distribution of preference distance follows an exponential pattern (Fig. 2.1C), panel C), suggesting that most students apply to closely related programs, while a smaller subset exhibits highly dissimilar preferences. The exponential nature of this distribution arises from the structural constraints of the higher education system, where degree program distances are not independent, leading to deviations from expected distributions under random selection. This implies that student choices are not random but structured by systemic constraints, reinforcing that preference alignment is a behavioral signal shaped by institutional and cognitive factors.

We hypothesize that students with more dissimilar preferences are less likely to persist in their enrolled program. First-year retention, defined as continued enrollment in the

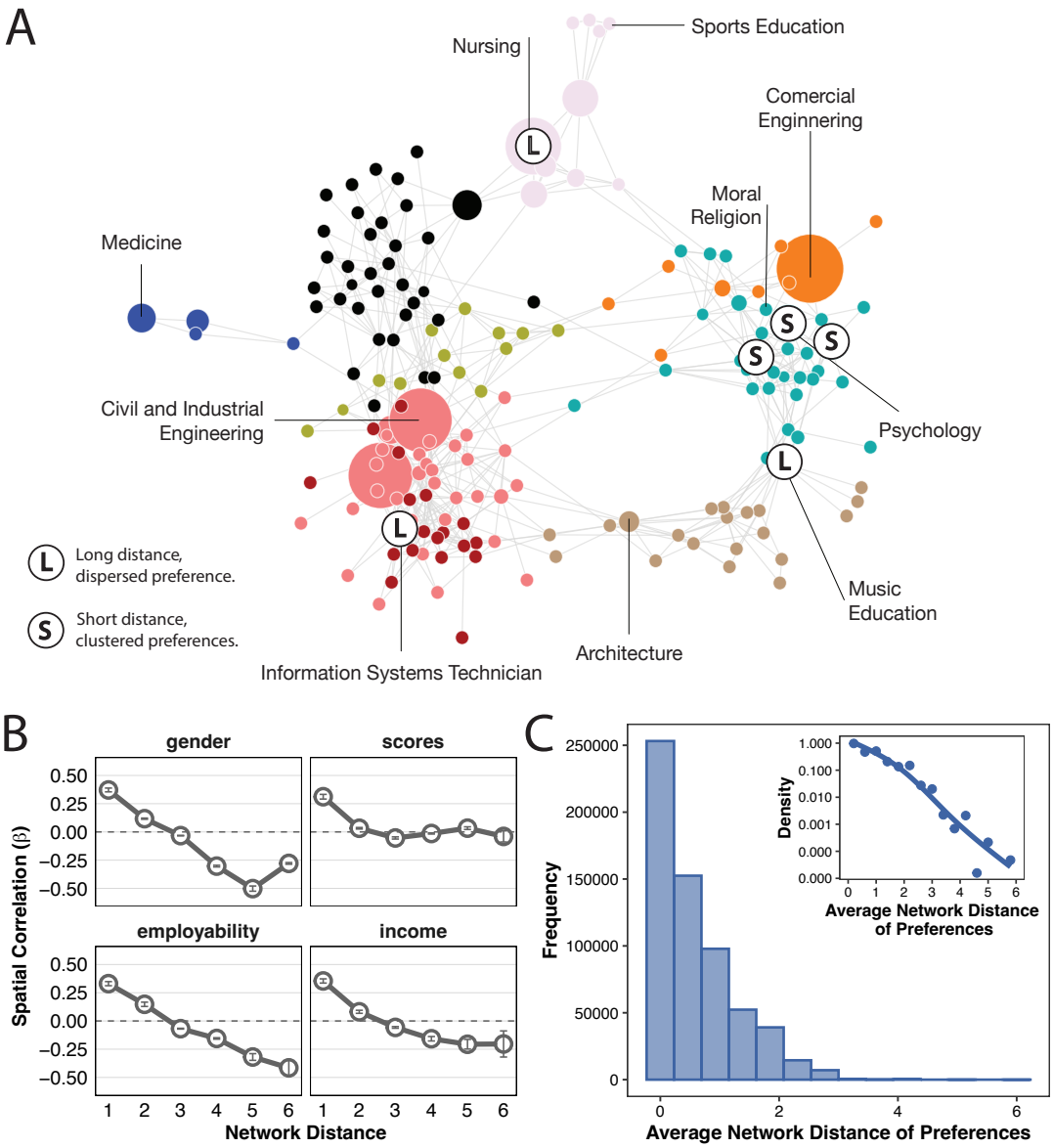

Figure 2.1: **Network structure of degree program preferences and distribution of preference distances.** **A**, Network of degree programs constructed from significant co-occurrence of programs in applicants' ranked preference lists. Nodes represent degree programs, and edges indicate statistically significant co-selection across applicants. Node size reflects the number of applicants listing the program as their first choice. Two regimes of preference alignment are illustrated: clustered preferences with low average network distance (L) and dispersed preferences with high average distance (H). Examples include Nursing, Music Education, and Information Systems Technician, which occupy distinct network regions. **B**, Spatial correlation between program-level attributes and network distance. Each panel shows how a distinct attribute–gender composition, application scores, employability, and expected income–varies as a function of average preference distance. Points represent mean correlations across programs and years at each network distance, with error bars indicating standard error. All attributes exhibit a systematic decay as distance increases, suggesting that closely related programs tend to align more strongly in their social and economic characteristics. **C**, Distribution of the Average Preference Distance across applicants, indicating that the majority select closely related programs, while a smaller subset display widely dispersed preferences. The inset shows the same distribution on a logarithmic scale, revealing a pattern consistent with exponential decay.

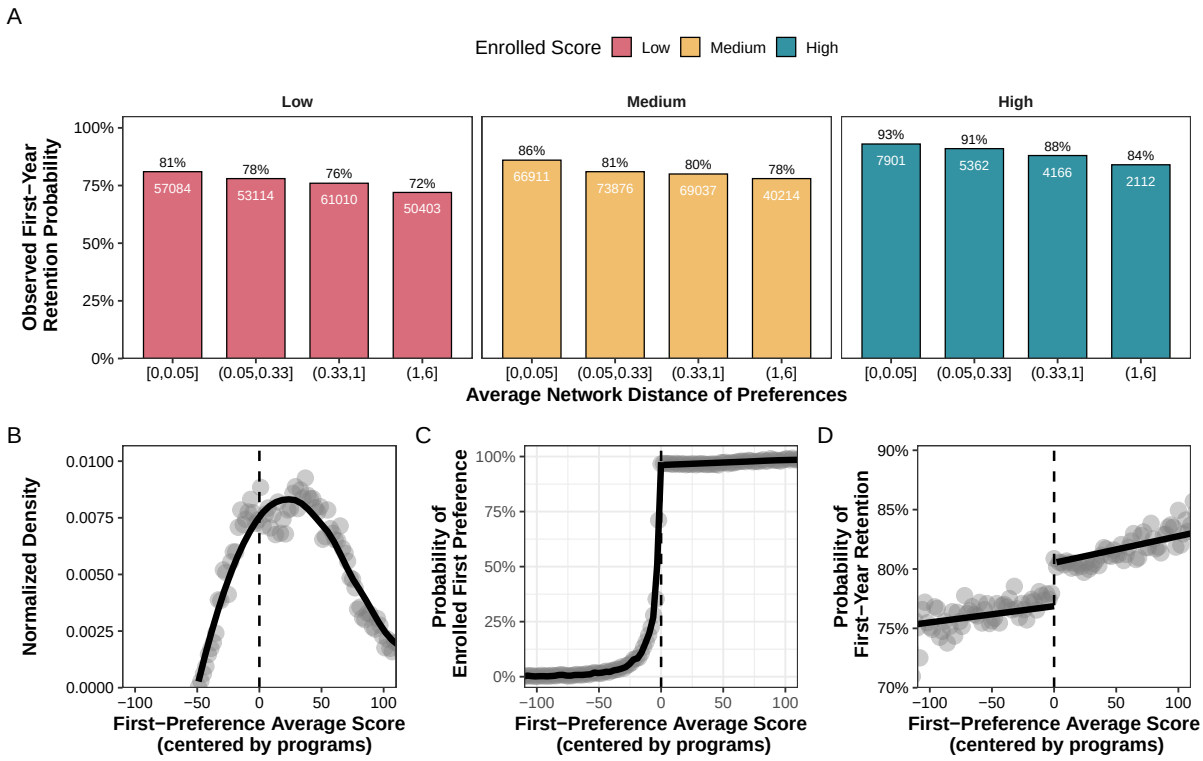

Figure 2.2: **A. Observed retention by preference alignment and student quality.** First-year retention probabilities are reported across bins of Average Preference Distance, stratified by tertiles of Enrolled Score. Retention declines monotonically with distance from preferred programs, particularly among students with lower scores. Each bar reports the retention rate and the total number of observations.**B–D. Empirical validation of the fuzzy regression discontinuity design. B.** The density of the running variable–defined as the difference between a student's score and the admission cutoff of their first-preference program–is continuous at the threshold, providing no evidence of strategic manipulation or sorting. **C.** The probability of enrolling in the first-choice program increases discretely at the cutoff, confirming the presence of a first-stage jump and validating the design as fuzzy rather than sharp. **D.** Retention also increases sharply at the threshold, indicating a causal effect of enrolling in one's top choice on first-year persistence. Together, these panels support the identification assumptions underlying the fuzzy RDD: continuity in potential outcomes and a nonzero first-stage discontinuity in treatment assignment.

same degree program for a second year, shows a decreasing trend with increasing preference distance (Fig. 2.2A). This inverse relationship holds across different score ranges, suggesting that both high- and low-scoring applicants experience the same negative effect of preference misalignment on retention. To formally assess the statistical effect of preference distance, we estimate logistic regression models where the dependent variable is first-year retention, taking a value of 1 if the student remains in their initial program and 0 otherwise.

Table A.1 (section A.1) presents the results of the fitted logistic models. We estimate three models: (i) a null model without preference distance, (ii) a model including preference distance as an independent variable, and (iii) an interaction model incorporating an interaction term between preference distance and the enrolled score. The results confirm that preference distance has a robust and significant negative effect on retention across all specifications. In the baseline model, for each one-unit increase in preference distance, the log-odds of retention decrease by 0.209 ($p < 0.01$). This effect remains consistent when

controlling for a comprehensive set of con-founders, including socioeconomic variables, school type, and parental education among many others.

The interaction model reveals that while students with higher scores have a greater overall likelihood of retention, they are not immune to the negative effects of preference distance. The interaction term between preference distance and enrolled score is significant ($p < 0.01$), indicating that even among high-scoring students, greater preference distance is associated with lower retention probabilities. If an applicant's preference distance is below 1, their retention probability exceeds 80%. However, for students with a preference distance above 4, retention probability falls to approximately 65%. Notably, for extreme values of preference distance, retention probability converges to 40% regardless of entry score, suggesting that when preference misalignment reaches a critical threshold, academic ability alone cannot offset its negative impact.

To investigate whether this relationship is causal, we exploit the quasi-experimental conditions created by the centralized admission system. The admission process in Chile assigns students to programs based on standardized test scores, generating a discontinuous threshold at each program's cutoff score. This setup allows for a regression discontinuity design (RDD), where students just above and just below the cutoff can be considered statistically equivalent in terms of both observed and unobserved characteristics, except for the opportunity to enroll in their first-choice program. The McCrary test results confirm that the distribution of scores remains continuous at the cutoff for high values, supporting the assumption of local randomization necessary for causal inference (Fig. 2.2B), see methods section 2.4 for more details.

We implement a fuzzy regression discontinuity design (FRD), which accounts for cases where students qualify for their first-choice program based on the cutoff but ultimately choose to enroll elsewhere, or vice versa, to estimate the local average treatment effect (LATE) of enrolling in the first-choice program on retention. This "fuzzy" variation in treatment assignment, where eligibility does not guarantee matriculation, is addressed using a two-stage least squares (2SLS) estimation. The first stage estimates the probability of enrollment in the first-choice program based on whether the student's score exceeded the cutoff, while the second stage uses this predicted probability as an instrument to estimate the effect of first-preference enrollment on retention.

Results indicate a significant discontinuity at the cutoff, confirming a robust first-stage relationship where eligibility increases the probability of top-choice enrollment; subsequently, our second-stage estimates demonstrate that this exogenous shift in enrollment significantly improves first-year retention probability for the subpopulation of compliers. To enhance the identification of the first stage and mitigate potential manipulation near the minimum assignment cutoffs, we employ a vectorial eligibility criterion (Z) as our instrumental variable. This instrument is defined by the intersection of meeting the program-specific cutoff and surpassing a high-performance threshold (Enrolled Score > 170):

$$Z_i = 1(\text{Running Variable}_i \geq 0 \cap \text{Enrolled Score}_i > 750) \tag{2.1}$$

The Z instrument isolates a source of exogenous variation that is less susceptible to

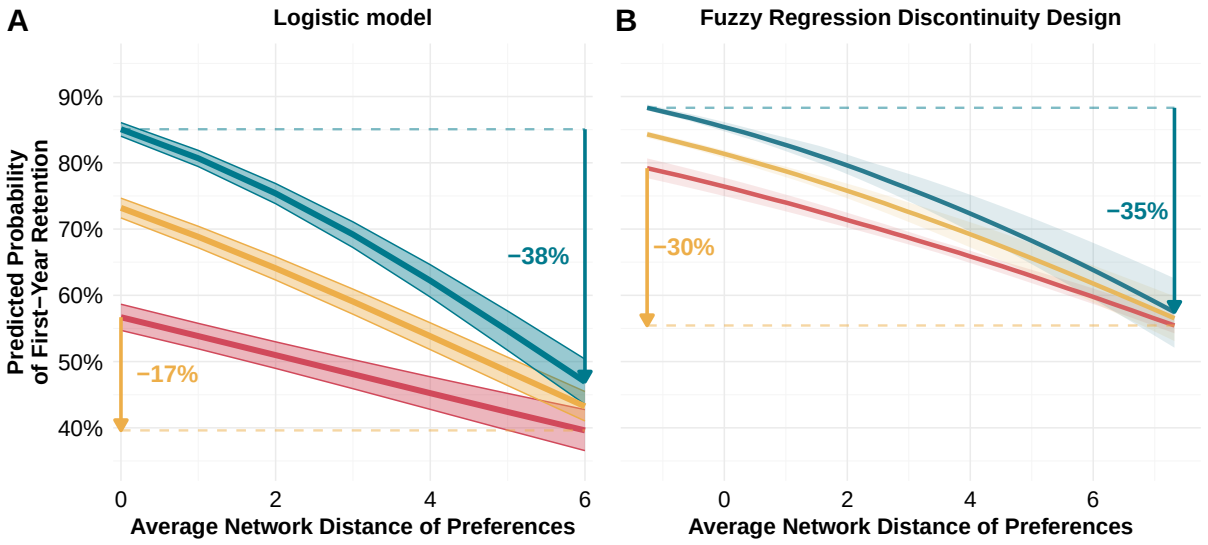

Figure 2.3: **Predicted probability of first-year retention by preference alignment and enrolled score.** Panels A and B show the predicted probability of retention as a function of the average network distance between the degree program students enrolled in and those they initially preferred, stratified by enrolled score (Low, Medium, High). Panel A presents results from a logistic regression, while Panel B estimates effects using a fuzzy regression discontinuity design; both including covariates. Shaded bands represent 95% confidence intervals. Arrows indicate the magnitude of the predicted decline in retention from the most aligned (lowest distance) to the most misaligned (highest distance) cases for high- and low-scoring students. Retention drops substantially for high-score students (-38% in Panel A, -35% in Panel B), but is more modest among low-score students (-17% and -30%, respectively). Dashed lines are included as visual guides only.

strategic sorting or gaming of the threshold. This approach ensures that the first-stage estimates are driven by students whose eligibility is determined by an external, multi-dimensional performance standard rather than marginal applicants near the cutoff. The strength of our first stage identification strategy is empirically confirmed by a robust First-Stage F-statistic of $1,202$. Following the criteria of Stock and Yogo (2005), this value comfortably exceeds the critical threshold of 16.38 for a 10% maximum IV bias, ensuring that our model is not subject to weak instrument concerns.

To validate the FRD assumptions, additionally to McCrary density test, that confirms the absence of systematic sorting, particularly for the individuals satisfying the vectorial eligibility criterion; also, our main specification bandwidth ($h = 100$) yields a LATE of 0.215 ($p < 0.01$). As shown in Table A.3 (A.1), our results demonstrate convergence across different bandwidth selection methods. Specifically, the bias-corrected CCT (MSE-optimal) selector ($h = 8.95$) yields a LATE of 0.222 ($p < 0.05$), which is nearly identical to our main specification. This consistency between the conservative CCT estimate and our CV-backed bandwidth confirms that our results are not an artifact of bandwidth selection but represent a stable causal relationship. These findings suggest that overcoming the initial enrollment threshold has a significant positive impact on long-term student retention.

Furthermore, the choice of bandwidth (h) serves as the structural boundary for the

Fuzzy Regression Discontinuity Design		
Second stage		
Dependent Variable:	First Year Retention (Logit)	
	Coefficient (β)	**Corrected SE**
Treatment Est. (LATE)	$\mathbf{0.308}^{***}$	(0.030)
Distance*	$\mathbf{-0.240}^{***}$	(0.009)
Running Variable	-0.0003	(0.0002)
Enrolled Score	$\mathbf{0.099}^{***}$	(0.013)
Constant	1.162^{***}	(0.022)
LATE Interactions ($\hat{X}$)		
Int Treatment × Distance	$\mathbf{0.074}^{***}$	(0.007)
Int Treatment × Running	0.001^{***}	(0.0003)
Int Treatment × Score	0.196^{***}	(0.018)
Int Distance × Score	$\mathbf{-0.024}^{***}$	(0.005)
Int Tratamiento × $D \times S$	$\mathbf{-0.012}^{***}$	(0.004)
First stage:		
Dependent Variable:	Enrollment in First Preference (Probit)	
Running	$\mathbf{0.044}^{***}$	(0.002)
Z Instrument	$\mathbf{0.807}^{***}$	(0.234)
Distance*	-0.008	(0.012)
Enrolled Score	$\mathbf{-0.153}^{***}$	(0.026)
Constant	0.177^{***}	(0.019)
Model Diagnostics		
First-Stage F-statistic	786.35	
Stock-Yogo Critical Value (10% bias)	16.38	
Observations: 447,173 (RDD Window [-100, 100])		

Note: Distance* and Score use standardized variables (std). Second Stage standard errors are based on Bootstrap inference (Cluster-Robust). First Stage standard errors are Cluster-Robust (vcovCL).
Significance: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 2.1: **Quasi-experimental design the case of Chile** Results reports from the Two-Stage Least Squares (2SLS) estimation strategy of a instrumental variable estimation using a Fuzzy Regression Discontinuity Design (FRD), to identify the Causal Effect of enrolling in one's top-ranked program on first-year retention. The first step (probit) estimates treatment assignment as a function of the running variable (score relative to program cutoff), the instrument (instrument), Average Preference Distance, and Enrolled Score. The instrument strongly predicts treatment assignment ($\beta = 0.807, p < 0.001$), validating the discontinuity as a source of exogenous variation. The second step (Logit) regresses first-year retention on the predicted treatment status ($\hat{X}$) and covariates. Receiving treatment increases retention by a significant amount ($\beta = 0.308, p < 0.001$), confirming a positive Local Average Treatment Effect (LATE). The model confirms that preference distance has a significant negative effect on retention ($\beta = -0.240, p < 0.001$).The positive coefficient for the Treatment × Distance interaction ($\beta = 0.074, p < 0.001$) suggests that the causal benefit of treatment increases as misalignment increases, operating as a source of causal mitigation. All reported standard errors are Cluster-Robust and Bootstrap-Corrected, and the model includes multiple interaction terms to control for heterogeneous treatment effects.

LATE. By utilizing a window supported by Cross-Validation and comparing it with the MSE-optimal selector ($h = 8.95$), we strictly delimit our causal inference to the local vicinity of the threshold. This ensures that the comparison is restricted to students who are statistically equivalent, thereby reinforcing that our estimates represent a local effect for compliers rather than a generalized average across the entire distribution of scores.

We further examine heterogeneous treatment effects by introducing preference distance as an interaction term in the FRD model. Figure 2.3 illustrates that the negative effect of preference distance on retention is significantly steeper for students who do not enroll in their first-choice program. For students with high preference distance (e.g., above 7.5 standard deviations), retention rates drop to approximately 30%, with a more pronounced decline among students enrolled in their non-preferred programs. The Second-Stage estimates reveal a significant average causal effect of first-preference enrollment on retention ($\beta = 0.308, p < 0.001$). The statistical validity of this complex GLM structure is supported by the model diagnostics. The substantial reduction from Null Deviance $(174,234)$ to Residual Deviance $(18,379)$ and the low AIC $(18,391)$ indicate that the inclusion of interaction terms and the predicted treatment significantly improves the explanatory power over a baseline linear specification

As shows in Table 2.1, when the treatment is absent, preference distance has a significant negative direct impact on retention ($\beta = -0.240, p < 0.001$), confirming that distance acts as a barrier to persistence. Interestingly, the interaction term between treatment and distance is positive and significant ($\beta = 0.074, p < 0.001$). This suggests that enrolling in a first-choice program acts as a risk-mitigation mechanism, the causal effect of being in one's top choice implies a reduction in the negative impact of preference misalignment. However, the significant triple interaction ($\beta_{T\times D\times S} = -0.012, p < 0.001$) further refines this, showing that the mitigating effect of preference alignment is slightly attenuated for students with high scores who also face long distances.

Our model evidences a critical vulnerability for students who fail to enroll in their first-choice program. For this group, the treatment effect is zero ($T = 0$), which triggers a cumulative negative impact on their retention probability. Specifically, these students face a dual penalty: the significant negative coefficient of preference distance ($\beta = -0.240, p < 0.001$) acts as a formidable structural barrier, while the negative interaction between distance and score ($\beta = -0.024, p < 0.001$) further erodes the persistence of high-achieving students who find themselves displaced from their intended preference clusters.

Without the mitigating shield provided by first-preference enrollment—namely the substantial LATE of 0.308 ($p < 0.001$) and the positive interaction of 0.074 ($p < 0.001$)—these individuals must bear the full weight of network misalignment. Consequently, the centralized system's failure to match a student with their top choice does not result in a merely neutral outcome; instead, it actively compounds structural risks. This synergy of negative pressures explains the observed collapse in retention rates, which drop to approximately 30% at high levels of preference distance for the subpopulation of non-compliers.

To assess the external validity of our findings, we replicate the analysis using data from Portugal, which has a similarly centralized admissions system but distinct institu-

tional and cultural features. The results are consistent across both contexts, reinforcing the robustness of the observed relationship between preference alignment and retention (See Supplementary Material A.1 for more details on the replication analysis). The instrumental variable validity tests confirm that the assignment variable satisfies the conditions for a valid instrument, and the similarity of results across countries suggests that preference alignment is a generalizable determinant of retention in higher education. Thus, preference alignment plays a crucial role in educational persistence, with greater alignment increasing the likelihood of retention. The heterogeneous treatment effects further highlight that misalignment reduce significantly the likelihood of successful institutional retention, particularly for students not enrolled in their first-choice program.

We provide further evidence that preference alignment acts as a behavioral signal of retention by assessing its predictive power. If preference alignment encodes cognitive decision-making under uncertainty, it should serve as an early indicator of student persistence. To test this hypothesis, we implemented a random forest classification model to predict first-year retention, incorporating preference distance as a key feature. By leveraging Bayesian optimization for hyperparameter tuning (Bischl et al., 2017; Wu et al., 2019), we assessed whether preference distance improves the predictive capacity of standard retention models.

The model achieved an overall accuracy of 82.2%, exhibiting a strong predictive capability (Fig. 2.4A). However, an imbalance emerged in the classification of non-retained students: while the model correctly predicted 144,563 retained students, it only correctly identified 7,767 cases of non-retention out of 37,930. This suggests that while preference distance provides significant explanatory power, retention is likely influenced by additional, unobserved factors.

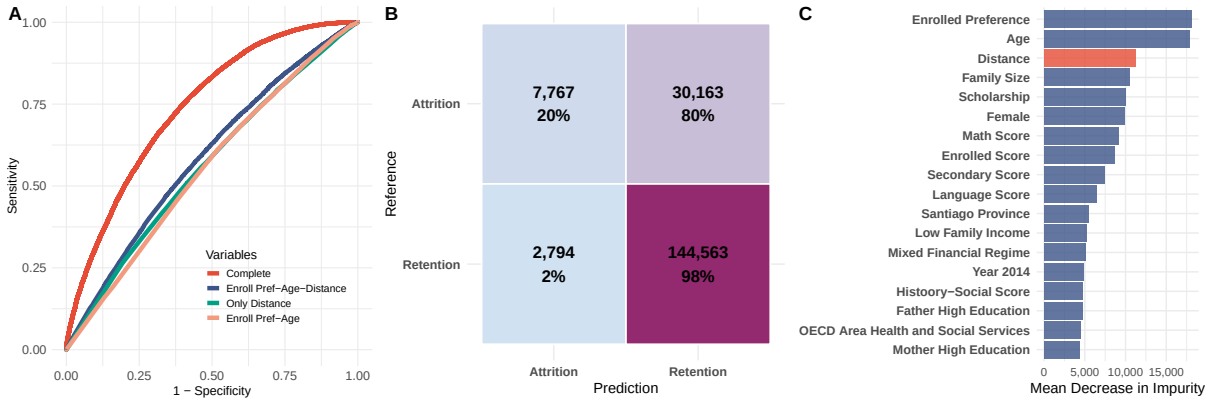

Figure 2.4: **Predictive modeling of first-year retention and feature importance.** **A** Receiver operating characteristic (ROC) curves comparing random forest models for predicting first-year retention using different subsets of predictors. The full model (Complete) achieves the highest discriminative performance, followed by models including Enrolled Preference (a proxy for degree program fixed effects), Age at enrollment, and Average Preference Distance. **B** Confusion matrix for the complete model reveals strong predictive accuracy for retention, but lower sensitivity for attrition detection, underscoring the asymmetry of prediction performance across outcomes. **C** Feature importance based on mean decrease in impurity shows that Enrolled Preference, Age, and the newly introduced Average Preference Network Distance are the top three predictors of first-year retention—outperforming traditional academic and socioeconomic variables. Distance is highlighted in orange to underscore its substantive contribution.

Feature importance analysis further underscored the relevance of preference alignment. According to impurity-based rankings (Gini index), preference distance ranked among the top three predictors of retention, surpassing traditional variables such as standardized entrance scores (2.4C). When using permutation-based feature importance, preference distance maintained a high ranking (Fig. 5D), reinforcing its relevance as a key behavioral signal driving retention.

To assess the independent predictive capacity of preference distance, we compared different model specifications. A model trained exclusively on preference distance exhibited meaningful classification power, comparable to age and enrollment preference (2.4A). However, its predictive power was enhanced when integrated with other academic and demographic variables, suggesting that preference alignment interacts with broader structural factors influencing student persistence.

2.4 Data and Methods

Our approach investigates the relationship between first-year retention and preference alignment by analyzing individual choices within the collective structure of higher education systems (Candia et al., 2019). Grounded in a social information perspective, we postulate that aggregate student behavior serves as a rich source of information, where historical patterns of co-application reveal a latent network of related degree programs. We quantify alignment not against an arbitrary optimal portfolio, but with respect to this collective knowledge that accumulates year after year.

2.4.1 Data Sources and Integration

We use data from the Department of Assessment, Measurement, and Educational Records (DEMRE) at the University of Chile, covering the period from 2005–2019. Our dataset integrates the following:

- Enrollment and Registration Data: To capture admissions outcomes and first-year retention rates.
- Application Preferences: To track portfolio preferences at the individual level.
- Scholarship and Socioeconomic Data: To account for financial and demographic factors influencing retention.
- Institutional and Program-Level Data: To contextualize students' choices and outcomes.

Focusing on data from 2012–2019, we build a comprehensive dataset to evaluate the interconnectedness of programs, student retention rates, and preference alignment. This dataset encompasses a network of careers, student retention rates, and individual student profiles.

By integrating retention data with co-application patterns, the methodology allowed for a comprehensive understanding of higher education dynamics in Chile, emphasizing both program connectivity and student persistence.

For the replication analysis in Portugal, to test the generalizability of our approach across different educational systems,, we use administrative data provided by the Direção-Geral do Ensino Superior (DGES), covering the higher education application cycles from 2008 to 2015. Similar to the Chilean case, this dataset contains the complete ranked lists of degree programs chosen by applicants, allowing for the construction of the Higher Education Space (HES) and the calculation of preference alignment metrics in a comparable centralized admission system (Candia et al., 2019).

2.4.2 Methods

We investigate whether a student whose preference set is misaligned with the collective academic portfolio exhibits lower first-year retention. This inquiry distinguishes between predictive correlation and causal impact, formalizing the degree of misalignment through a topological network approach. Our objective is to determine to what extent this relationship is driven by causal mechanisms—identified via quasi-experimental designs—and the degree of inferential and predictive power provided by the observed structural patterns. Specifically, we examine the hypothesized negative relationship between holding a preference set distant from the collective "social optimum" and the probability of academic persistence. To achieve this, we develop an integrated methodological framework that combines generalized linear models, fuzzy regression discontinuity designs, and random forest algorithms.

Construction of the Higher Education Space

To construct the higher education space, we used application data from Chile's centralized admissions system spanning 2012–2019. Each degree program is represented as a node, and links between programs are established based on co-occurrence in students' preference sets. Following the methodology of (Candia et al., 2019), we computed the ϕ-correlation between pairs of programs and retained only positive, statistically significant correlations (p-value ≤ 0.05). The network was binarized, such that links represent the presence or absence of a significant correlation. Degree programs were classified using generic degree codes, which group together similar academic offerings across different universities.

Higher education applicants navigate a complex decision-making process that involves selecting degree programs under uncertainty. To examine the structure of applicants' preferences, we constructed a network of degree programs based on application data from 2012 to 2019, comprising 156 nodes (Fig. 2.1A) that correspond to generic degree programs used in the national higher education system. This network connects pairs of degree programs according to their co-occurrence in applicants' preference sets, allowing us to infer their similarity.

For each year, we computed co-occurrence counts M_{ij} between degree programs i and j as the number of applicants listing both programs in their preference set. Let $M_i = \sum_j M_{ij}$ denote the total number of co-occurrences involving program i, and let Z be the total number of co-occurrences in the dataset. We quantified pairwise association using the ϕ-correlation

$$\phi_{ij} = \frac{M_{ij}Z - M_iM_j}{\sqrt{M_iM_j(Z - M_i)(Z - M_j)}}.$$

Negative associations were discarded by setting links only for pairs with $\phi_{ij} > 0$. To ensure that links were not due to chance, we applied two one-tailed tests (p-value ≤ 0.05). First, we compared the observed ϕ_{ij} against a null distribution obtained from $N = 1000$ randomized networks generated by shuffling applicants' preference lists within each year while preserving (i) the number of preferences per applicant and (ii) the yearly frequency of each degree program; the p-value was computed as the upper-tail probability of observing a value at least as large as ϕ_{ij} under this null model. Second, we tested whether ϕ_{ij} differs from zero using the statistic

$$t_{ij} = \frac{\phi_{ij}\sqrt{D - 2}}{\sqrt{1 - \phi_{ij}^2}}, \qquad D = \max(M_i, M_j),$$

and retained links only when the corresponding p-value was ≤ 0.05. Finally, we removed self-links ($i = j$) and binarized the resulting network by keeping an unweighted edge whenever the pair (i, j) satisfied the criteria above (Candia et al., 2019).

To assess whether the structure of student preferences remains rigid or fluid over time, we conducted a longitudinal analysis of the network. We constructed independent annual networks for the 2012–2020 period and tracked the "lifespan" of each edge.

We define the "birth" of an edge at time t_0 when its proximity metric exceeds a significance threshold ($\phi_{ij} \geq 0.05$), and its "death" at time t_k when ϕ_{ij} falls below this threshold or the edge disappears entirely.

The probability of link persistence at time t was estimated using the non-parametric Kaplan-Meier estimator (Kaplan & Meier, 1958):

$$\hat{S}(t) = \prod_{t_i < t} \left(1 - \frac{d_i}{n_i}\right) \tag{2.2}$$

where d_i denotes the number of edges that "died" at time i, and n_i represents the number of edges at risk of disappearing. This approach allows us to quantify the structural inertia of the system while appropriately accounting for right-censored data—specifically, those edges that remained active until the end of the observation period.

Preference Distance Metric

The network structure allows us to define a topological distance measure between degree programs. The dissimilarity score of an individual applicant is calculated as:

$$C_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \sum_{j=i+1}^{N_k} d_{ij} \tag{2.3}$$

where N_k represents the number of degree programs in the application list P_k, and d_{ij} is the shortest path distance between degree programs i and j in the network. Sensitivity analyses were performed considering 3, 5, and 10 preferences, with results demonstrating robustness across different specifications.

Regression Models for Retention Analysis

We estimate logistic regression models to examine the relationship between preference distance and first-year retention. The dependent variable is retention, defined as continued enrollment in the same program in the second year (coded as 1) versus non-retention (coded as 0). Independent variables include preference distance, admission test scores, gender, family income, school type, parental education, and scholarship status. Control variables are included to account for known predictors of retention and persistence in higher education.

Three models are estimated:

1. A null model excluding preference distance.
2. A model incorporating preference distance as an independent variable.
3. An interaction model with an interaction term between preference distance and enrolled score.

Regression Discontinuity Design (RDD)

To establish causality, we employ a fuzzy regression discontinuity design (FRD), leveraging the cutoff scores that determine program admissions. This design acknowledges that treatment assignment is not solely determined by the threshold.

The causal parameter τ_{LATE} is identified through the ratio-form Wald estimand, which formalizes the relationship between the discontinuity in the outcome and the discontinuity in treatment assignment. This estimator is defined as:

$$\tau_{LATE} = \frac{\lim_{s \downarrow 0} E[R|S = s] - \lim_{s \uparrow 0} E[R|S = s]}{\lim_{s \downarrow 0} E[T|S = s] - \lim_{s \uparrow 0} E[T|S = s]} \tag{2.4}$$

Where the components of the equation are defined as follows:

- R: Represents the outcome variable, **First-Year Retention**.

- T: Represents the treatment variable, **Enrollment in the First-Choice Program**.
- S: Denotes the **Running Variable**, defined as the student's admission score relative to the program-specific cutoff.
- $s \downarrow 0$: The limit approaching the threshold from the right, representing eligible applicants.
- $s \uparrow 0$: The limit approaching the threshold from the left, representing ineligible applicants.

The numerator of the equation captures the reduced-form jump in retention at the threshold, while the denominator measures the discontinuity in the probability of treatment—referred to as the first-stage effect. In a fuzzy design, the denominator is typically less than one, as eligibility does not guarantee enrollment. Consequently, the estimator "scales" the observed shift in retention to account for non-perfect compliance. This mathematical structure ensures that the resulting coefficient represents the average causal effect specifically for the subpopulation of complier—those whose enrollment decision was strictly triggered by crossing the eligibility threshold—effectively excluding the noise from never-takers and maintaining the validity of the local randomization.

The estimation is implemented using a Two-Stage Least Squares (2SLS) framework within a Generalized Linear Model (GLM) to account for the binary nature of both enrollment and retention. The first stage of the model estimates the probability of enrollment in the first preference based on test scores, while the second stage uses this probability as an instrument to estimate its effect on retention.

Model Specification, Identification Strategy, and Monotonicity

- *First-stage equation:*

$$T_i = \alpha_0 + \beta_1 Z_i + \beta_2 (S_i > 0) + \mathbf{X}'\delta + \varepsilon_1 \tag{2.5}$$

where T_i is the binary indicator for enrollment, S_i is the running variable, and Z_i is the instrument defined as being above the cutoff conditional on eligibility criteria.

The estimation of causal effects in the centralized admission systems of Chile and Portugal presents significant identification challenges. Since students apply to programs after receiving their entrance exam scores, a potential self-selection issue arises. Applicants may strategically limit their choices based on prior knowledge of previous cutoffs or opt for waiting lists, potentially leading to non-random sorting near the threshold.

To address these endogeneity concerns and the "Fuzzy" nature of enrollment—where eligibility does not guarantee matriculation—we implement a robust Instrumental Variable (IV) framework within a Regression Discontinuity Design (RDD). We utilize what we term a vectorial eligibility criterion, where the treatment assignment is identified through a composite instrumental variable. This binary instrument (Z_i) identifies high-scoring (> 750) students who exceed the admission cutoff for any of their applied programs. This accounts for potential selection biases related to application strategies.

$$Z_i = 1(\text{Running Variable}_i \geq 0 \cap \text{Enrolled Score}_i > HighScore) \tag{2.6}$$

It is important to clarify that our focus is not restricted to a high-scoring subpopulation; rather, we utilize this high-performance eligibility criterion as our instrumental variable in the first stage. This strategy specifically addresses the link between self-selection and potential manipulation. While self-selection might lead to strategic sorting near the minimum cutoff for marginal applicants, the subpopulation satisfying our vectorial instrument—high-scoring individuals—faces fewer constraints and lacks the incentive for "gaming" the threshold. By utilizing this cleaner discontinuity, we ensure that the estimated Local Average Treatment Effect (LATE) is not biased by the strategic behavior of marginal applicants.

A critical requirement for the empirical identification is the monotonicity assumption ($T^1 \geq T^0$), which implies the absence of defiers. In the context of centralized admission systems in Chile and Portugal, this assumption is guaranteed by the institutional architecture. Meeting the program-specific cutoff ($Z = 1$) acts as a technical eligibility requirement that only increases or maintains the probability of enrollment; it is logically impossible for a student to be induced to enroll by failing to meet the cutoff, nor would meeting it discourage an otherwise willing applicant. While some students may qualify and choose not to enroll (never-taker), the system's rules ensure that the instrument moves the treatment probability in a single direction.

A potential challenge to the monotonicity assumption in centralized systems is the fact that applicants are not blind to the selection process. Students receive their test scores prior to application and have access to historical cutoff scores (c_{t-1}), allowing them to act strategically. This creates a risk of strategic sorting, where students only apply to programs where they perceive a high probability of compliance ($T = 1$). Specifically, this behavior may lead to endogenous sorting (D. S. Lee & Lemieux, 2010; McCrary, 2008; Urquiola & Verhoogen, 2009), a situation where the running variable is no longer locally exogenous because applicants precisely manipulate their application sets to guarantee a specific treatment status.

If students with higher unobserved motivation or better information are the ones who strategically align their preferences with the threshold, the resulting LATE would be biased by this self-selection (Urquiola & Verhoogen, 2009). However, following the framework of (Bertanha & Moreira, 2020), we argue that as long as there is an irreducible degree of uncertainty regarding the current year's equilibrium cutoff (c_t), the local randomization property near the threshold remains valid. In this context, endogenous sorting might affect the composition of compliers, but it does not violate monotonicity unless the strategic behavior induces defiance—a logical impossibility in a system where higher scores strictly expand rather than contract the choice set (Choi & Lee, 2023).

- *General Second-stage equation:*

$$R_i = \alpha_0 + \tau_{LATE}\hat{T}_i + f(S_i) + \mathbf{X}'\gamma + \varepsilon_2 \tag{2.7}$$

where R_i represents first-year retention and $\hat{T}_i$ is the predicted treatment.

Bandwidth Selection

To determine the optimal estimation window, we conducted a comprehensive sensitivity analysis comparing data-driven bandwidth selectors—including the Mean Squared Error (MSE) optimal bandwidth (Calonico et al., 2014) and Imbens-Kalyanaraman (IK) methods—against cross-validation procedures. While MSE-optimal methods suggested narrower local windows ($h \approx 9 - 37$), we selected a bandwidth of $h = 100$ for the main specification. This choice ensures sufficient statistical power to estimate complex interaction terms while maintaining the robustness of the LATE estimate, as confirmed by our sensitivity checks across multiple bandwidths (see Supplementary Material). To validate this selection, we contrast our results with the Cross-Validation (CV) criteria and the bias-corrected MSE-optimal selector proposed by Calonico, Cattaneo, and Titiunik (CCT).

The resulting estimate is interpreted strictly as a Local Average Treatment Effect (LATE). This estimand is localized to compliers whose enrollment was induced by the exogenous threshold, and its validity is restricted to the specific bandwidth h around the program cutoffs.

Heterogeneous Treatment Effects and Robust Inference

The final model incorporates interaction terms between the instrumented treatment ($\hat{T}$) and preference distance to evaluate heterogeneous treatment effects:

$$R_i = \cdots + \gamma_2 C_i + \gamma_3(\hat{T}_i \times C_i) + \varepsilon_2 \tag{2.8}$$

where C_i represents preference distance.

Crucially, to ensure valid inference given the use of a predicted regressor ($\hat{T}_i$) in a non-linear second-stage model (Logit), standard analytical errors are insufficient. We computed Cluster-Robust Standard Errors using a Non-Parametric Bootstrap procedure with $N = 500$ iterations. This method rigorously corrects for the generated regressor bias and accounts for the correlation of errors clustered at the Year level ($G = \text{Year}$), providing robust t-statistics for hypothesis testing of the LATE and its interactions.

The final two-stage least squares (2SLS) model incorporates an interaction term with preference distance to evaluate heterogeneous treatment effects.

$$R_i = \alpha_0 + \beta_1(S > 0) + \beta_2(S - 0) + \gamma_1 \hat{T}_i + \gamma_2 C_i + \gamma_3 \hat{T}_i C_i + \varepsilon_2 \tag{2.9}$$

where C_i represents preference distance, allowing us to capture how the treatment effect varies by preference misalignment.

Predictive modeling of student retention

To evaluate the predictive power of preference alignment in forecasting first-year retention, we implemented a supervised machine learning approach using a random forest classifi-

cation model. This method was selected due to its ability to handle high-dimensional data, capture complex nonlinear relationships, and provide robust estimates of variable importance. The model was trained and tested on a large-scale dataset containing student application and retention records, ensuring that the findings were generalizable within the studied educational system.

The random forest classifier was optimized through a hyperparameter tuning process that relied on Bayesian optimization, a probabilistic approach that efficiently searches the hyperparameter space to maximize model performance. The key hyperparameters included the number of decision trees, the maximum depth of each tree, and the minimum number of samples required for node splitting. Five-fold cross-validation was used to prevent overfitting, where the dataset was partitioned into five subsets, iteratively training on four while validating on the remaining fold. The final model was trained on 80% of the dataset, with the remaining 20% reserved for evaluation.

Feature selection and importance assessment were conducted to isolate the role of preference distance in predicting retention. Two complementary methods were employed to determine the relative contribution of each predictor. First, impurity-based feature importance, measured using the Gini index, quantified the reduction in uncertainty attributable to each variable. Second, the Boruta algorithm, a permutation-based method, assessed the robustness of each predictor by iteratively introducing random shadow variables and comparing their contributions. Preference distance was analyzed in relation to traditional predictors, including age, standardized entrance scores, socioeconomic background, and gender.

The performance of the predictive model was evaluated using several classification metrics. Overall accuracy was computed as the proportion of correctly classified cases in the test set. A confusion matrix provided a detailed breakdown of classification performance, distinguishing between correctly and incorrectly predicted retention and attrition cases. Additionally, the receiver operating characteristic (ROC) curve was used to assess the predictive power of the model across different classification thresholds by plotting sensitivity against specificity.

To ensure a rigorous assessment of the independent predictive power of preference alignment, multiple model specifications were tested. A baseline model incorporating all available predictors was compared against models that excluded preference distance, models that relied exclusively on preference distance, and models that included key interaction terms, such as the interaction between enrollment preference and preference distance, as well as the interaction between age and preference distance. This approach allowed for a nuanced understanding of the extent to which preference distance contributes to retention outcomes beyond traditional academic and demographic predictors.

The dataset comprised 617,622 student records spanning multiple application cycles. To prepare the data for modeling, categorical variables such as parental education and income group were converted into binary indicators using one-hot encoding. Continuous variables, including entrance scores, were standardized to ensure comparability across students. Missing values were imputed using a k-nearest neighbors (k-NN) approach, which preserved the underlying distributional properties of the data.

The results of the predictive analysis indicated that preference distance was a significant determinant of retention. The final model achieved an overall accuracy of 82.2% on the test set; however, the confusion matrix revealed that while retention cases were well-predicted, the model struggled to accurately classify instances of attrition. Preference distance emerged as one of the three most important predictors of retention, ranking above entrance scores in importance. A model trained solely on preference distance exhibited meaningful classification power, comparable to age and enrollment preference, reinforcing the idea that preference alignment serves as a strong behavioral signal of retention. These methodological choices ensured that the estimation of preference alignment's predictive utility was robust, offering insights into its role in educational decision-making beyond conventional indicators.

External Validity through Replication

Since the main focus is on the method and the demonstration of the validity of the measure of relatedness between degree programs to indicate educational outcomes, we extend the analysis to the Portuguese case, which has a similar admission system. Being able to replicate the results allows us to give strength to policy recommendations and subsequent lines of research.

We test the generalizability of our approach by replicating all the methods and analysis using in the case of chilean data, that implies generates for Portugal: the higher education spacial, the preference distance measure, inference models, quasi-experimental models, and machine learnig models, with the same assumptions and robust checks.

2.5 Discussion

This study shows that preference alignment, measured through the distance between selected degree programs, is a key determinant of student retention in higher education. By leveraging large-scale data and causal inference techniques, we provide robust evidence that students with aligned degree preferences exhibit significantly higher first-year retention rates than those selecting dissimilar programs. These findings contribute to ongoing discussions on decision-making under uncertainty (Alchian, 1950; Manski, 1991; Manski, 2004; Tversky & Kahneman, 1974) and highlight the role of informational constraints and institutional structures in shaping student outcomes (Hastings et al., 2018; Larroucau et al., 2024; Ye, 2024).

Our results indicate that first-choice enrollment improves retention only when preference alignment is high, reinforcing the importance of considering preference structures in higher education decision-making. While previous research has extensively documented the role of academic preparation, socioeconomic background, and institutional resources in student persistence (Bourdieu, 1977; Mitchall & Jeager, 2018; O'Neill et al., 2011; Stewart et al., 2015; Voelkle & Sander, 2008), our findings highlight that preference misalignment itself constitutes a distinct barrier to student retention. This suggests that retention is not solely a function of student ability or institutional quality, but also of how applicants structure their academic choices under uncertainty (Barnett, 2007; Flum & Blustein,

2000; Perez et al., 2014b). The relationship between preference alignment and retention highlights the critical role that program-related information, resources, and institutional factors play in shaping students' decision-making. In centralized higher education systems, students rank degree programs with imperfect information about acceptance probabilities, career trajectories, and program quality ((Hastings et al., 2018; Ye, 2024)). This uncertainty may lead some applicants to select dissimilar programs, either as a hedging strategy (Chade & Smith, 2006) against rejection or due to limited vocational guidance (Flum & Blustein, 2000; Moore et al., 2020; Perez et al., 2014b; Santelices et al., 2025). Our findings suggest that such application behaviors may significantly compromise academic persistence, particularly among students who fail to enroll in their first-choice program.

From an institutional perspective, the network-based preference distance metric reveals a systematic pattern in application behaviors, where a majority of students select closely related programs while a smaller subset applies to highly dissimilar degrees. This distribution suggests that while some students enter higher education with clearly defined academic goals, others apply with broader uncertainty, potentially diminishing their probability of continued enrollment. The exponential distribution of preference distance further underscores that misaligned preferences are not randomly distributed but systematically shaped by the higher education application process. These findings align with prior research on higher education decision-making under uncertainty (Kelly, 2001; Pucciarelli & Kaplan, 2016), which suggests that students often rely on their social system of support and on incomplete information to construct their preference lists. This process may reinforce inequities in access to high-quality educational pathways (Belasco & Trivette, 2015; Kuh et al., 2006; J. Smith et al., 2013), particularly for first-generation or low-income students (Mitchall & Jeager, 2018; Network, 2011). Expanding access to reliable information on degree programs and career trajectories (Hastings et al., 2018; Ye, 2024) may help mitigate these disparities.

By implementing a fuzzy regression discontinuity design (FRD), we provide causal evidence that first-choice enrollment only improves retention when degree program distance is low. This finding suggests that first-choice enrollment is not a universal predictor of persistence but rather moderated by preference alignment. Students admitted to a first-choice program with highly misaligned preferences exhibit retention rates similar to those placed in non-preferred programs, reinforcing that preference alignment is an essential component of educational success. Our findings build upon prior research on selection problems in higher education (Larroucau & Rios, 2022; Solis, 2017) and show that students' enrollment decisions do not operate in isolation but are shaped by both individual preferences and systemic constraints. Furthermore, the heterogeneous treatment effect analysis reveals that students with high preference distance face an additional retention penalty when they do not enroll in their first-choice program. This effect is particularly strong for students with preference distances exceeding 7.5 standard deviations, where retention probabilities drop to approximately 30%. These findings suggest that preference misalignment exacerbates the consequences of non-preferred enrollment, making it a key vulnerability factor for student persistence.

Additionally, these results align with broader discussions on heterogeneous causal effects in policy evaluations (Angrist, 2004; Athey & Imbens, 2016; Athey & Imbens, 2017, 2019), reinforcing that not all students experience the same retention benefits from first-

choice enrollment. These insights highlight the need for targeted interventions that account for individual preference structures and application behaviors. Replication in Portugal confirms that preference alignment is a robust predictor of retention across distinct institutional and cultural contexts, underpinning its generalizability. While Chile and Portugal share centralized admissions frameworks, their higher education landscapes differ in terms of institutional structures and socioeconomic environments. The consistency of our results across both countries suggests that preference alignment is not country-specific ,but rather a fundamental component of student decision-making and critical for persistence.

These findings have direct implications for higher education policy. If preference alignment is a structural predictor of retention, early interventions could mitigate its negative effects. Policies aimed at enhancing student decision-making, improving access to degree program information, and providing targeted vocational guidance could encourage informed-based program applications, bolstering system-wide retention rates. Informational interventions, such as those previously tested in Chile (Hastings et al., 2018) and China (Ye, 2024), have shown that improving access to program-specific information enhances student-program alignment and retention. The school system could implement preference alignment metrics as early-warning indicators in student advising planning. Machine learning models, as shown in our predictive analysis, can identify at-risk students based on preference distance, enabling secondary and tertiary educational institutions to offer personalized interventions (Bettinger & Baker, 2014; Nieuwoudt & Pedler, 2023). Additionally, structured exploratory programs, modular curricula, and flexible academic pathways could provide students with greater adaptability, ensuring that early misalignment does not result in long-term attrition (Tinto, 1993; Zeidenberg et al., 2007).

A natural extension of this research would involve longitudinal studies tracking student trajectories over time to examine how preference alignment evolves, and whether changes in academic interests impact retention. Measuring students' program alignment through high school (9th-12th grade) may offer insights into their vocational growth and alignment upon graduation and future university enrollment. Understanding how early academic experiences influence students' choices can inform interventions to tackle students' knowledge gaps. Additionally, controlled interventions, such as early orientation workshops, could provide direct insights into strategies that effectively support students with dissimilar preferences. Since preference alignment reflects not only individual decision-making but also broader institutional dynamics, further investigation into how universities can adapt admissions policies and academic pathways to support students with varying degrees of alignment would be valuable.

Beyond the scope of individual institutions, future research could also examine the role of preference misalignment in broader labor market outcomes, exploring whether early misalignment affects long-term career trajectories. Investigating the role of social capital and access to information in shaping student preferences could further illuminate the mechanisms behind preference misalignment. By integrating insights from decision science, educational policy, and network analysis, future studies can continue to refine our understanding of how preference structures shape long-term academic and professional outcomes.

Addressing preference misalignment is not merely about maximizing retention num-

bers; it is about creating educational environments where students can thrive in programs that align with their interests, abilities, and long-term goals. By prioritizing equity, adaptability, and informed decision-making, higher education systems can foster more inclusive academic pathways and enhance student success while optimizing public investment in education.

Chapter 3

A Computational Social Approach to Decoding Sex Differences in Autism

3.1 Abstract

The increasing prevalence of autism spectrum condition (ASC), contrasts with persistent diagnostic challenges in females. In addition to sex-based behavioral and biological differences, the propensity for autistic females to mask autistic characteristics creates significant barriers to accurate diagnosis, particularly in women without intellectual disability (ID). Understanding clinical-characteristic patterns in ASC presentation is indispensable for developing appropriate supports and improving autism identification across sexes. To address this issue, our study analyzed sex-specific differences (and overlaps) in ASC characteristics in a sample of 97 autistic girls and boys aged 6-17 years without ID, combining network methodologies with complementary computational approaches. Our results showed a significantly later age of diagnosis in girls, averaging two years later than in boys. Furthermore, significant differences were observed in standardized scales assessing eye contact, language use, verbal reasoning, quantitative reasoning, processing speed, and overall cognitive proficiency. We constructed a bipartite network linking individuals and clinical characteristics, and constructed sex-specific three-layer multilayer networks (ADOS-2, cognitive indices, and co-occurring psychological and medical conditions), testing structural invariance between girls' and boys' networks with the Network Comparison Test (1000 permutations) and quantifying cross-domain coupling via inter-layer strength and node versatility. Analysis of these network architectures revealed distinct organizational patterns of clinical characteristics and co-occurring conditions between sexes. Specifically, we identified sex-divergent topological predictors of diagnostic delay: while phenotypic centrality in clinical characteristics facilitated earlier diagnosis in boys, it was associated with delayed diagnosis in girls. Additionally, the integration of co-occurring medical and psychological conditions exhibited distinct structural configurations, suggesting sex-specific organizational configurations of these co-occurring conditions. Specifically, autistic girls exhibit a complex and diffuse integration characterized by high inter-layer coupling between cognitive performance and psychological co-occurring conditions; in contrast, boys show a more modular and targeted architecture, where medical co-occurrences act as

structural anchors to specific cognitive nodes. In addition, cluster analysis incorporating ADOS-2 variables and cognitive performance revealed at least two sex-differentiated subgroups, which further validated the existence of distinct phenotypic profiles. These findings contributed to an understanding of autism as a complex system of interrelated characteristics rather than a sum of a number of independent traits that are independent expressions of a single latent condition. We discussed the importance of these types of approaches to achieve a more refined understanding of autism, which can guide the development of sex-sensitive diagnosis and support alongside relevant public health policies.

Keywords: Autism Spectrum Condition, Neurodevelopmental Disorders, Sex differences in Autism, Computational Social Science, Symptomatological Networks.

3.2 Introduction

Sex differences in prevalence have been reported across several psychiatric and neurodevelopmental conditions, such as depression (Caqueo-Urízar et al., 2020; Jiménez-Molina et al., 2021), anorexia nervosa (Gaete P & López C, 2020; Neale & Hudson, 2020), and attention-deficit/hyperactivity disorder (Chung et al., 2019). This is also the case for Autism Spectrum Disorder, a neurodevelopmental condition characterized by persistent challenges in social communication and interaction across multiple contexts, and by restrictive and preferred routines and in-depth interests (American Psychiatric Association, 2013b).

Sex Differences and Overlaps in Autism

Autism Spectrum Condition (ASC) has been consistently reported as more prevalent in males than females, with a historically accepted male-to-female ratio of approximately 4:1 (Baio et al., 2018; Dellapiazza et al., 2022; Fombonne, 2009; Loomes et al., 2017a; Roman-Urrestarazu & van Kessel, 2022; Zeidan et al., 2022). However, this estimate has been increasingly questioned, as research indicates that various factors may influence the male-to-female ratio—particularly a subtler presentation of autistic characteristics often seen in girls without intellectual disability, especially in social functioning (Dworzynski et al., 2012; Ratto et al., 2018; Wood-Downie et al., 2021). Underdiagnosis in girls may result from both biological and social variables, including the limited capacity of diagnostic tools to account for gender-related variations in autistic characteristics, as well as the presence of compensatory strategies such as social camouflage (Begeer et al., 2013; Hull et al., 2017; M. C. Lai et al., 2017). These misdiagnoses or delays in diagnosis are clinically significant, given their association with increased risks of anxiety, depression, and reduced quality of life, particularly in autistic females (J. V. Smith et al., 2024).

From a clinical perspective, social functioning features have been broadly documented in autism, including challenges in emotional reciprocity, nonverbal communication, and the ability to develop and maintain relationships (Caruana et al., 2015; Hamilton, 2009; Jones & Klin, 2013; Kana et al., 2014; May & Kana, 2020; Mundy, 2017; Soto-Icaza et al., 2019). Preferred routines and in-depth interests has also been described as a core feature of autism, typically presenting as rhythmic movements, echolalia, ritualistic routines, or

intense focus on specific topics or objects (**edition2013diagnostic**; Lord et al., 2015). Notably, evidence has shown that these characteristics present distinct patterns depending on sex.

Recent efforts to characterize autistic females have revealed gender-related differences in both social expression and interests. In this regard, it has been described that girls tend to engage in more socially normative interests or autism-related themes (e.g., animals, arts and crafts), while boys often focus on stereotypically masculine topics such as video games or mechanical systems (Brown et al., 2024; Edwards et al., 2024). However, studies have not only identified differences, but also substantial overlap in many interest categories across sexes, particularly in areas such as TV/movies and music (Brown et al., 2024). Furthermore, evidence has found that autistic boys tend to display less male-typical play than non-autistic boys, showing preferences for female-typical activities or playmates (To & Kung, 2025). In contrast, autistic girls did not show significant differences in play patterns compared to non-autistic girls, although some tendencies toward more male-typical preferences were observed in areas such as toy and activity choices.

Similarly, a recent study found that both autistic girls and non-autistic girls tend to smile more and with greater prototypicality (i.e., higher correlation between lip corner pulling and cheek raising) than boys (Zampella et al., 2024). Overall, autistic boys and girls tend to smile less and their smiles are less prototypical compared to non-autistic peers, with no significant differences between sex and diagnosis interactions. Interestingly, in autistic individuals, the quality of social interactions appears to be more closely linked to smile expressivity in males than in females, possibly suggesting different social dynamics across sexes. Moreover, findings showed that autistic boys were more likely to display repetitive movements such as hand-flapping and intense interest in mechanical parts, while autistic girls showed more subtle focused interests, including collecting items like stickers or pens, and exhibited greater distress with minor changes (Edwards et al., 2024; Hiller et al., 2016). Evidence in preschool-aged autistic girls showed that they tended to present fewer repetitive behaviors than boys (Hiller et al., 2016), and that their strong personal interests were less mechanically oriented, often reflecting either socially typical themes or unique areas of interest.

Importantly, studies have argued that under-recognition of autism in girls could be related with the phenomenon of social camouflage, in which autistic individuals, particularly females with higher cognitive abilities, consciously or unconsciously mask social-communication difficulties (Buttelmann et al., 2009; Cook et al., 2021; S. J. Howe et al., 2023; Jedrzejewska & Dewey, 2022; M. C. Lai et al., 2017; Montagut et al., 2018; Russo, 2018). This may involve learning social scripts, imitating peers, using rehearsed phrases, and forcing eye contact, which—while improving superficial social presentation—often comes at a significant emotional and cognitive cost (S. J. Howe et al., 2023; Russo, 2018). Furthermore, caregivers frequently report that girls exhibit a heightened desire to please others, and that their masking behaviors are often misinterpreted as characteristics of non-autistic behavior (Hiller et al., 2016).

In addition, evidence has shown that elements associated with cognitive performance and language development in boys and girls influence the differences observed between genders (Dellapiazza et al., 2022; Y. J. Howe et al., 2015; J. V. Smith et al., 2024).

More specifically, autistic girls who exhibited language development with phrases and fluent speech showed similar or better social and adaptive skills than boys, while those who were minimally verbal showed cognitive, adaptive, and social skills similar to boys. These findings support evidence that girls require greater challenges in multiple domains to meet diagnostic thresholds (Dworzynski et al., 2012). Taking this evidence together, and considering the complexities arising from variations in how autistic characteristics are expressed, diagnostic constraint, and the risk of under-recognition of ASC in girls, we analyzed sex-specific differences in autistic characteristics among autistic children and adolescents aged 6 to 17 years without intellectual disability. For this purpose, we used a computational social science approach to examine features across sexes.

Network Analysis and Autistic Characteristics

Network analysis offers a rigorous framework for studying autism as a complex system of interrelated characteristics. Instead of assuming that observable traits are independent expressions of a single latent condition, network models describe them as components that influence one another through direct and indirect relationships (Borsboom, 2017; Epskamp et al., 2018). For instance, bipartite network models can be a useful tool to capture such complexity, by connecting individuals to the features they express. In the case of our data, bipartite network models provide a way to link behavioral descriptors and cognitive indicators without collapsing these relationships into total scores. This structure makes it possible to identify subgroups that share expressive configurations rather than diagnostic labels. Such an approach is especially relevant for understanding the diagnostic challenges in autistic girls (Dworzynski et al., 2012; Ratto et al., 2018; Wood-Downie et al., 2021). From a network perspective, these challenges, partially arising from subtle or atypically integrated patterns of behavior in autistic girls, can be understood as configurations characterized by nodes with strong connections across domains but weaker links within specific clusters. They represent atypical yet internally coherent structures that remain partially invisible to traditional diagnostic systems.

Autism is multidimensional, encompassing behavioral, cognitive, and medical aspects. Multilayer network frameworks (Kivelä et al., 2014) extend the understanding of this multidimensional complexity by representing it across multiple interacting levels, including clinical characteristics (e.g., ADOS-2 items), cognitive indicators (e.g., WISC-V subscales), and co-occurring conditions (e.g., anxiety symptomalogy). Each layer represents an internal organization, while connections between layers capture how features interact across domains. Within this structure, pathways of diagnostic overlap can become visible, helping to identify intermediate profiles (Hus & Segal, 2021; Muris & Ollendick, 2021). Accordingly, evidence on the overlap between attention-deficit/hyperactivity disorder (ADHD) and autism characteristics based on parent reports from a community sample of children aged 6–17 years shows limited cross-construct associations. In a joint ADHD–autism network analysis (Farhat et al., 2022), only about 10% of possible cross-construct connections were different from zero, revealing limited overall interconnectivity. However, specific ADHD characteristics expressed in social contexts, such as "talks excessively" or "does not wait for their turn," showed moderate to strong associations with autistic features. These bridges can help to explain diagnostic challenges when symptoms occur at the intersection of distinct clinical domains.The presence of overlapping features and bridge nodes suggests that different conditions may share functional pathways even

when categorical boundaries remain distinct (Afzali et al., 2017; Casanova et al., 2020; Farhat et al., 2022; Isvoranu et al., 2022; Sommers et al., 2025; Van Heijst et al., 2020).

In addition, evidence in autistic and non-autistic children and adolescents shows that peripheral domains, such as sleep or executive function, play an active role in maintaining the network structure (K. S. Lee et al., 2024; Sommers et al., 2025). Taken together, these findings suggest that autism may be conceptualized as a dynamically coupled system in which the configuration of connections shapes the prominence and recognizability of distinct features.

To address these complexities, the present study combines network methodologies with complementary computational approaches. In addition to bipartite and multilayer network models, we use regression analyses to estimate associations between network metrics and diagnostic timing, and clustering algorithms to identify subgroups characterized by distinct relational configurations. These complementary tools provide different but compatible insights: networks reveal structural organization, regressions quantify predictive relationships, and clustering captures emergent group patterns. Similar multimethod frameworks have been successfully applied in research on co-occurring conditions and neurodevelopment (Borsboom & Cramer, 2013; Waldren et al., 2024). This integrative strategy links the architecture of relationships to measurable outcomes, bridging the methodological gap between network science and the statistical models of computational social science, underscoring its relevance for understanding complex behavioral systems.

Studying these relations directly allows network research to move beyond the question of which features differ between groups and instead investigate how the structure of relationships among characteristics varies across populations. This perspective aligns with the concept of recognitional justice (Fraser, 2000). Diagnostic systems privilege certain structural configurations as recognizable while neglecting others. Network analysis makes this mechanism empirically visible by showing how gendered and cultural patterns of interconnection shape who is identified, when, and under what profile (Diemer et al., 2022). In this way, network science contributes not only new analytical tools but also a theoretical lens for understanding the diagnosis of autism as a matter of relational visibility and institutional recognition.

3.3 Data and Methods

Ethics statement

All methods and the experimental protocol received approval from the Ethics Committee of Universidad del Desarrollo, adhering to the principles outlined in the Declaration of Helsinki as well as the Local Ethical Guidelines for Research Involving Human Subjects. The study ensured informed consent was obtained from both the parents or legal guardians of the children and the children provided their assent to participate voluntarily. The responses provided by boys and girls were treated with strict confidentiality, and their identities were anonymized using alphanumeric codes.

3.3.1 Participants

The sample consisted of 97 autistic children and adolescents (47 girls, 50 boys) aged 6 to 17 years (Mean age = 11.18 ± 3.07). No statistically significant differences were found in age between groups (Wilcoxon test, $W = 1025$, $p = 0.28$, effect size $r = 0.31$). All children were born at term and were native Spanish speakers.

Participants were recruited through private clinicians, therapists, foundations, and private healthcare facilities. The inclusion criteria required participants to have a confirmed clinical diagnosis of Autism Spectrum Condition (ASC) according to DSM-5 criteria (American Psychiatric Association, 2013a) and an Intelligence Quotient (IQ) ≥ 70 points. Exclusion criteria included participants with features of possible genetic disorders, confirmed diagnosis of genetic disorders, or auditory and/or visual sensory deficits.

3.3.2 Behavioral and Cognitive Measures

Cognitive and psychological assessment of the participants considers the use of standardized scales in order to define the behavioral domains. This complete evaluation considered:

Cognitive assessment. The IQ was estimated in order to homogenize the sample. Therefore, participants aged from 6 years 0 months to 16 years 11 months were evaluated using the Wechsler Intelligence Scale for Children – Fifth Edition (WISC-V), standardized edition for Chile (Rodríguez et al., 2019) Participants aged older than 17 years were evaluated using the Wechsler Intelligence Scale for Adults – Fourth Edition (WAIS-IV), standardized edition for Chile (Rosas et al., 2014).

Social and communicative assessment. Participants of all chronological ages were evaluated with the Autism Diagnostic Observation Scale - 2 (ADOS-2) (Lord et al., 2015). This scale is a standardized, semi-structured assessment of communication, social interaction, and play or imaginative use of materials for individuals suspected of having an autism spectrum disorder. In the present study, Module 3 or Module 4 was used depending on each participant's age and language level. Each of these modules is made up of a set of activities that provide standardized contexts where the evaluator can observe or not the presence of certain social and communicative behaviors relevant to the diagnosis of ASC. The main scores that can be obtained are: 0 points (behavior not observed), 1 point (mildly atypical behavior, less pronounced than a score of 2), 2 points (clearly atypical behavior).

In addition, parents answered the SCQ Social Communication Questionnaire (Rutter, M., Bailey, A., & Lord, 2003). This questionnaire can be completed by parents or caregivers in approximately 10 minutes. The SCQ is made up of 40 items and delivers a total score plus three possible additional scores (Social Interaction Problems, Communication Difficulties, and Restricted, Repetitive, and Stereotyped Behavior). The questionnaire is presented in two forms: Form A, which refers to the entire past life of the subject, and form B, which must be answered considering the behavior presented during the last 3 months. According to the objectives set out in this research, the application of form A

was considered.

Finally, a semi-structured interview was carried out with the parents in order to identify potential psychological and co-occurring medical conditions.

3.3.3 Data Analysis

Our research adopts a computational social science approach to uncover interpretable and testable patterns in characteristic configurations across autistic boys and girls. This interdisciplinary framework allows us to integrate psychological, cognitive, and medical data through three complementary modeling strategies, aiming to identify structure within the heterogeneity of autism in both sexes.

Variable Operationalization

Building upon the assessments described in the previous section, we transformed the raw clinical measures into a structured dataset suitable for computational analysis. This process involved standardizing scores across different age-appropriate scales and selecting shared items to ensure comparability across the sample. The following domains and their respective variables were operationalized to serve as the features for our multilayer network and clustering models:

Cognitive Performance

In the case of the **WISC-V** (Wechsler Intelligence Scale for Children - Fifth Edition), standardized for the Chilean population (Rodríguez et al., 2019), we included the standardized scores of the 12 subscales required to obtain the "Full Scale IQ", the five "Primary Index Scales" (Verbal Comprehension, Visual Spatial, Fluid Reasoning, Working Memory, and Processing Speed), and the five "Secondary Index Scales" (Quantitative Reasoning, Auditory Working Memory, Nonverbal, General Ability, and Cognitive Proficiency). We also considered the scores of the Index Scores, and the standardized scores of each of the 12 subtests individually: Similarities, Vocabulary, Block Design, Visual Puzzles, Matrix Reasoning, Figure Weights, Digit Span, Picture Span, Coding, Symbol Search, Letter-Number Sequencing, and Arithmetic.

In the case of the **WAIS-IV** (Wechsler Adult Intelligence Scale - Fourth Edition), standardized for Chile (Rosas et al., 2014), we selected the same standardized subtest scores included in the WISC-V, except the *Picture Span* score, as this subtest is not part of the Chilean WAIS-IV version.

Autistic Characteristics

- **ADOS-2:** The Autism Diagnostic Observation Schedule, Second Edition (Modules 3 and 4), was used to assess social-communication and repetitive behaviors. To ensure comparability across modules, we selected algorithm items shared by both

modules. Specifically, the selected items correspond to the dimensions of the diagnostic algorithm "Communication, Social Interaction, and Restricted and Repetitive Behaviors". The shared algorithm items selected were: *Conversation; Descriptive, Conventional, and Instrumental gestures; Unusual eye contact; Facial expressions directed to others; Reciprocal social communication; Excessive interest in unusual, highly specific topics; Compulsions or rituals; Reporting events; Quality of social overtures; Quality of social response; Stereotyped/Idiosyncratic words or phrases; Unusual sensory interest; Shared enjoyment in interaction; Overall quality of rapport; and Hand and finger/Other complex mannerisms.*

- **SCQ:** The Social Communication Questionnaire (Form A - Lifetime) provided parent-reported developmental history. Specifically, item 35 ("Make-believe play") was used as a proxy for symbolic play deficits.

Co-occurring Conditions A semi-structured parental interview was conducted to identify co-occurring conditions. For the purpose of network analysis, these were stratified into two distinct domains:

- **Psychological Conditions:** Grouped into categories including *Mood Symptomatology* (depressive, hypomania, mania); *Anxiety Symptomatology* (social anxiety, anxious expectation, phobias); *Behavioral Symptomatology* (oppositional defiant, heteroaggression, emotional dysregulation); *Attentional Symptomatology* (inattention, distractibility); *Obsessive Compulsive and Related Symptomatology* (tics, Body-Focused Repetitive Behaviors, rituals); *Language Symptomatology* (articulation errors, expressive/receptive disorders); *Psychotic Symptomatology*; *Sleep Symptomatology*; *Eating Symptomatology*; *Suicide & Self-Harm Symptomatology*; and *Distress Symptomatology* (irritability, guilt, anguish).
- **Medical/Physical Conditions:** Including *Gastrointestinal Symptoms* (constipation, bloating, nausea); *Allergic Symptoms* (rhinitis, insect-sting allergy); *Neurological/Psychomotor Issues* (hypotonia, developmental coordination disorder); *Lactose Intolerance*; *Psoriasis*; *Asthma*; *Otolaryngological Alterations*; *Obesity*; *Thyroid Disorders*; and *Phimosis.*

Statistical Evaluation of Sex Differences

To generate the descriptive profile of sex differences, we did not rely solely on mean comparisons. Instead, we fitted generalized linear models to estimate the effect of sex on each variable while handling the different data types appropriately. For continuous variables (e.g., IQ standardized scores), we used linear regression models ('lm'). For binary variables (e.g., presence/absence of co-ocurrency conditions), we used logistic regression models ('glm' with binomial family). The resulting coefficients (Beta for linear, Log-Odds for logistic) were used to visualize the magnitude and direction of the differences between boys and girls.

Network-Based Prediction of Diagnostic Timing

We examined sex-based differences in the age of diagnosis, operationalizing diagnostic timing in two ways: a linear regression predicting age at diagnosis (in years) and a logistic regression predicting the probability of Late Diagnosis. We defined late diagnosis as occurring at or after 8 years of age, a threshold that distinguishes early-childhood from middle-childhood identification and captures diagnoses occurring well beyond national averages (≈ 4.8 years in Chile) (García et al., 2021) and international medians (Diemer & Gerstein, 2024; Maenner et al., 2023).

We constructed an undirected bipartite network $G = (U, V, E)$, where the set of nodes U represents the autistic individuals ($N = 97$) and the set of nodes V represents the clinical characteristics. In this framework, each edge in E represents the presence of a specific clinical characteristic in a given individual. To handle the heterogeneous nature of the data, we applied a specific discretization strategy before network assembly:

- **Cognitive nodes (IQ):** Standardized scores were discretized into three categorical levels based on standard deviations: "Low" (< 7), "Avg" ($7 - 13$), and "High" (> 13). This transformation allows the network to capture specific cognitive profiles (e.g., a child connected to *Vocabulary_ High* vs. *Vocabulary_ Low*).
- **Autism-related behavior (ADOS-2) nodes:** We adopted a granular approach in which each item was decomposed into independent nodes based on its severity levels (e.g., **Eye Contact equal to 0, Eye Contact equal to 2**). The edges are binary, connecting each individual to the specific node that represents their observed score for that item. This structure allows the network to capture the full distribution of behaviors, including the absence of autistic-related characteristics, as distinct relational states.
- **Co-ocurring nodes:** Binary variables were included as nodes only when the condition was present.

Following the graph construction, the topological structure was formally encoded into an incidence matrix $\mathbf{M}$ of dimensions $N_{characteristics} \times N_{patients}$. To ensure computational consistency across metrics, the graph object was converted and transposed to strictly align rows with clinical characteristics (i) and columns with individuals (j). Thus, the matrix entry $M_{ij} = 1$ denotes that patient j presents characteristic i, while $M_{ij} = 0$ indicates its absence. This algebraic representation served as the foundation for calculating vector-based metrics such as Entropy and Divisiveness.

Topological Metrics:

To ensure conceptual validity, metrics were calculated directly on the bipartite topology or its spectral properties, avoiding information loss associated with one-mode projections. For example, the following metrics were computed for each child node u_i (to see all the metrics calculated go to B tabla B.1):

- **Betweenness Centrality:** A key predictor in our regression models. It measures the extent to which a child's profile acts as a "bridge" connecting disparate autistic-related characteristics clusters. High betweenness indicates a profile that integrates otherwise disconnected phenotypic regions.
- **Hub Score (HITS):** We employed the HITS algorithm (Hyperlink-Induced Topic Search) to calculate the Hub Score (Kleinberg, 1999). In this clinical context, this metric serves as a robust measure of 'Clinical Load', identifying children who connect to the most authoritative nodes in the network. In network science, the concept of authority refers to nodes that are central due to their strong relational consistency within the system. Here, these nodes represent autism-related characteristics that are observed, specifically those where behavior is mildly atypical (score 1) or clearly atypical (score 2). By excluding the absence of symptoms (score 0, in which behavior is not observed), the score effectively quantifies the extent of present autism-related characteristics for each child. This structural ranking approach follows established methods in complex health data for identifying key clinical associations (Ghulam et al., 2020).
- **Divisiveness:** Quantifies the fragmentation of an individual's profile. High divisiveness indicates a profile dispersed across multiple autism-related characteristics clusters, lacking a dominant focus.
- **Domain Entropy:** A measure of phenotypic diffusion calculated using Shannon entropy over the distribution of a child's connections across the three clinical domains (cognitive, autism-related characteristics, medical). High entropy indicates a "noisy" or diffuse profile spanning all domains.
- **Harmonic Centrality:** Used as a robust alternative to Closeness Centrality for potentially disconnected graphs. It measures the average shortest distance from a child to all other nodes (autism-related characteristics), reflecting phenotypic accessibility.

Prior to their inclusion in the regression models, all topological metrics were standardized using Z-score transformation ($z = \frac{x-\mu}{\sigma}$). This procedure rescales the variables to have a mean of 0 and a standard deviation of 1, ensuring that the regression coefficients (β) are comparable across metrics with different units (e.g., entropy vs. degree) and facilitating the interpretation of interaction effects with sex.

Robustness and Validation of Topological Metrics: To ensure the reliability of the topological predictors, we conducted extensive robustness checks on the interaction between sex and network metrics. Model stability was confirmed through collinearity diagnostics (variance inflation factor, VIF < 6) and heteroskedasticity-robust standard errors. We employed bootstrapped confidence intervals ($B = 2000$) and leave-one-out estimations to verify that the observed patterns were not driven by influential observations. Furthermore, nonlinearity tests using natural splines confirmed that the linear specification provided the optimal fit ($F = 0.18$, $p = 0.94$).

Multilayer Architecture and Structural Invariance

Building upon the bipartite analysis, we constructed sex-specific multilayer networks to examine how characteristic domains interact structurally. The framework consisted of three interconnected layers: (1) ADOS-2 behavioral characteristics, (2) Cognitive Performance (WISC-V/WAIS-IV indices), and (3) Co-occurring Conditions (Psychological and Medical).

Edge Definition:

Connections were defined using Spearman correlations to accommodate the ordinal nature of clinical scores.

- Intra-layer edges: Represent correlations between variables within the same domain.
- Inter-layer edges: Represent correlations between variables from different domains.

A hard threshold of $|r| > 0.25$ was applied to retain only robust associations.

Metrics:

We computed node versatility, defined as the strength of a node's inter-layer connections (the sum of weights of edges connecting to a different layer), and the participation coefficient, which quantifies the diversity of a node's connections across the three layers.

We analyzed the structural invariance between sex-specific networks using the network comparison Test (NCT) with 1000 permutations, implementing a custom estimator to respect the specific construction parameters of our networks. Finally, we computed inter-layer strength to quantify connectivity between domains and node versatility to identify bridge nodes.

Identification of Latent Phenotypic Profiles

To explore whether distinct latent profiles could be identified beyond sex-based comparisons, we conducted an unsupervised cluster analysis. The variable sex was excluded from the clustering procedure to allow for the independent identification of latent profiles.

Given the mix of categorical and continuous variables, we computed a dissimilarity matrix using Gower's distance. The clustering algorithm selected was Partitioning Around Medoids (PAM) due to its robustness to outliers. The optimal number of clusters ($k = 4$) was determined by maximizing the Average Silhouette Width, testing solutions ranging from $k = 2$ to $k = 10$.

3.4 Results

Given the multidimensional nature of Autism Spectrum Condition (ASC), we adopt a progressive analytical approach to disentangle sex differences in diagnostic patterns. We begin by examining specific variations across clinical domains, including Autism Diagnostic Observation Schedule, Second Edition, (ADOS-2), intelligence quotient (IQ), and co-occurring conditions. Moving beyond individual variables, we then integrate topological metrics derived from a bipartite network to enhance the prediction of diagnostic age, revealing how the structure organization of clinical characteristics modulates diagnostic identification. To capture variability beyond sex, we subsequently present an exploratory cluster analysis identifying distinct clinical-cognitive profiles. Finally, we map the complex architecture of the clinical presentation by analyzing the structural interplay between cognitive, behavioral, and medical layers within a multilayer network.

3.4.1 Sex differences are embedded in specific sub-domains

The sample characterization revealed that the participants showed an ADOS-2 mean total score of 11.03 ± 3.39 (SD). Their mean total IQ was 103.65 ± 15.16 (SD), while the mean Social Communication Questionnaire (SCQ-A) total score was 14.84 ± 6.62 (SD). Notably, no statistically significant differences were found between autistic girls and boys across these measures (e.g., ADOS-2: Wilcoxon test $p = 0.12$, effect size $r = 0.16$; IQ: Wilcoxon test $p = 0.75$, effect size $r = 0.03$; SCQ-A: Wilcoxon test $p = 0.13$, effect size $r = 0.34$) (Table 3.1).

It is worth noting that although most variables did not show significant deviations from a normal distribution between autistic girls and boys, the ADOS-2 scores for the autistic girls' group showed evidence of non-normality (Shapiro-Wilk test: $W = 0.903$, $p < 0.001$). In contrast, scores in the boys' group did not deviate significantly from normality ($W = 0.97$, $p = 0.30$). Therefore, the Wilcoxon rank-sum test was selected as a conservative and robust non-parametric alternative to the t-test. This approach accounts for the non-normal distribution observed in the autistic girls' group while ensuring the validity of the overall comparison, as the Wilcoxon test does not assume normality and is less sensitive to outliers or distributional irregularities.

Variable	**Girls** (N = 47)	**Boys** (N = 50)	***t*-test** (*p*-value)	**Effect Size** (*r*) [95% CI]
Mean age (years)	11.6 ± 3.35	10.8 ± 2.76	0.28	0.11 [0.00, 0.31]
ADOS-2 score	10.5 ± 2.87	11.5 ± 3.80	0.12	0.16 [0.00, 0.35]
IQ score	103 ± 14.3	104 ± 16.1	0.75	0.03 [0.00, 0.24]
SCQ-A score	13.6 ± 6.58	16.0 ± 6.54	0.13	0.34 [0.00, 0.34]

Abbreviations: SD = standard deviation; ADOS-2 = Autism Diagnostic Observation Schedule, Second Edition; IQ = Wechsler Intelligence Scale for Children – Fifth Edition (WISC-V) and Wechsler Intelligence Scale for Adults – Fourth Edition (WAIS-IV); SCQ-A = Social Communication Questionnaire version A.

Table 3.1: Demographic and Clinical Characteristics by Sex.

To identify sex-based differences within the latent structures emerging from the sample, an exploratory analysis combining categorical and continuous behavioral-cognitive variables was formed. Specifically, we constructed contingency tables crossing sex with ADOS-2 algorithm item scores (Modules 3 and shared Module 3/4 items). For each item, we used a(χ^2) test of independence when expected cell counts were adequate, and Fisher's exact test otherwise. In addition, sex differences in standardized IQ subtest and index scores (WISC-V/WAIS-IV) were evaluated using Wilcoxon rank-sum tests, while differences in variability were assessed using both an F-test for homogeneity of variances and the Fligner-Killeen test. The latter being more robust to deviations from normality and appropriate for comparing spread in non-normally distributed cognitive data. When tests were significant, distributions were further examined descriptively by comparing category proportions.

The exploratory analysis revealed significant sex-specific patterns across cognitive, behavioral, and clinical domains, as summarized in Figure 3.1. Regarding cognitive performance (Panel A), autistic girls exhibited a significant difference in Coding and Symbol Search, while boys showed higher performance in Figure Weights. In terms of autism-related characteristics (Panel B), boys presented more pronounced atypicalities in eye contact, shared enjoyment and conversation, whereas girls showed higher scores in stereotyped language. Finally, the clinical landscape of co-occurring conditions (Panel C) highlighted a significant male preponderance in OCD and related symptomatology.

Autistic girls scored significantly lower than boys on the Visual Puzzles subtest, with group differences observed in both central tendency and variance (Wilcoxon test, $p < 0.05$; F-test, $p < 0.05$). A similar pattern of variance differences was found for the Vocabulary subtest (F-test, $p < 0.05$), although mean differences were not statistically significant. Additionally, autistic girls showed higher performance on Coding and Symbol Search subtests, with statistically significant differences in mean scores in both subtests (Wilcoxon test, $p < 0.05$).

The analysis of ADOS-2 items revealed significant sex-specific distributions that point toward subtler atypicalities in girls. With respect of Item B1 (Unusual eye contact), chi-

squared (χ^2) test indicated that girls were significantly more likely to receive a score of 0 (not observed) compared to boys (girls = 0.59, boys = 0.26, $p = 0.0039$), while boys more often received a score of 2 (clearly present). In Item B2 (Facial expressions directed toward the examiner), a global difference in distribution emerged (Fisher's exact test, $p < 0.05$), driven by a higher proportion of boys receiving a score of 0 ($p = 0.040$) and a higher concentration of girls exhibiting mildly atypical behavior (score 1: girls = 0.90, boys = 0.75).. Finally, Item A4 (Stereotyped or idiosyncratic use of words or phrases) also showed significant sex differences: boys were more likely to receive a score of 0 ($p = 0.025$), whereas girls were more likely to receive a score of 1 ($p = 0.013$), suggesting more subtle language-related atypicalities among girls. No differences were found for score 2. These findings suggest the presence of latent behavioral-cognitive structures that differ by sex within the autistic sample.

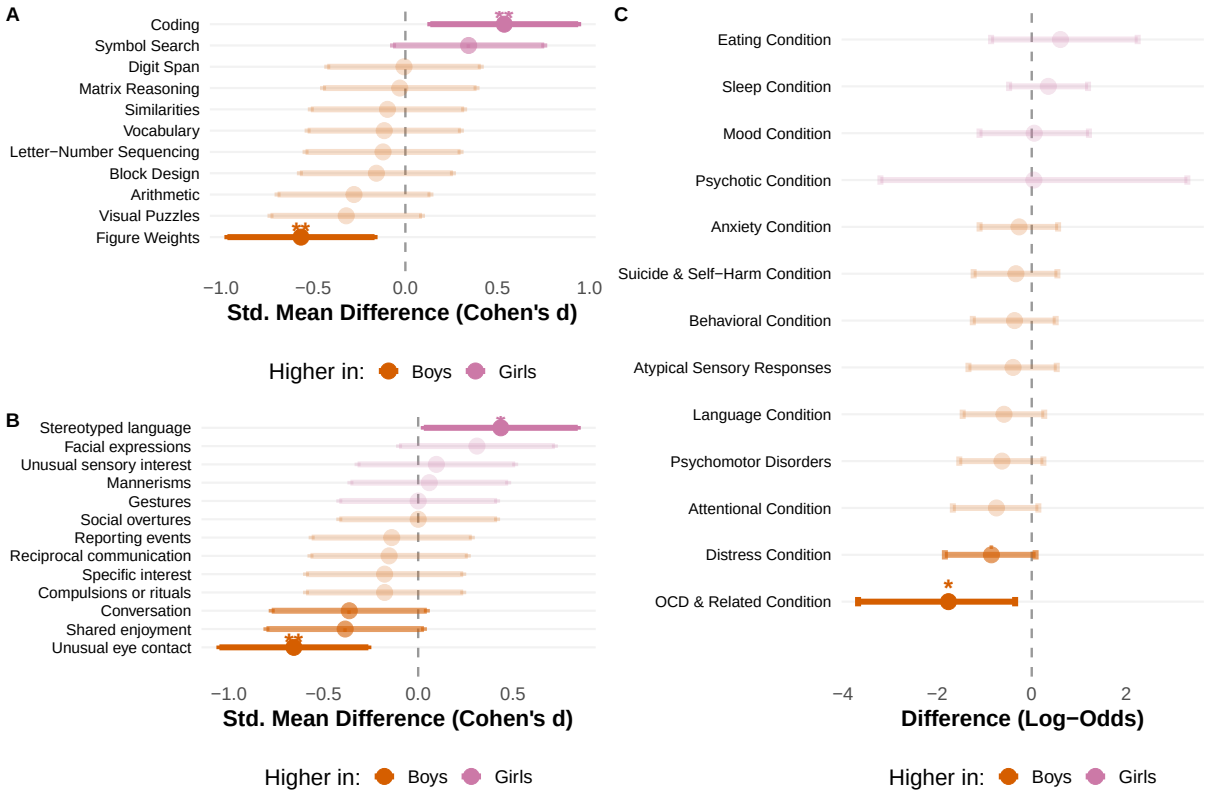

Figure 3.1: **Sex-Specific Differences across Clinical Domains.** **A** Cognitive Profile (IQ): Standardized Mean Differences (Cohen's d) in cognitive subtests. Girls (purple) exhibit higher performance in processing speed tasks (Coding, Symbol Search), while boys (orange) show higher scores in Figure Weights). (Rodríguez et al., 2019). **B** Autism Profile (ADOS-2): Standardized differences in algorithm items. Sex differences were observed across different components of social communication and interaction, with boys showing higher scores in the items Unusual Eye Contact, Conversation, and Shared Enjoyment, and girls showing higher scores in Stereotyped Language. **C** Co-occurring Conditions: Log-Odds ratios for binary medical and psychological conditions. The plot highlights a male preponderance in OCD & Related Symtomatology. Error bars represent 95% confidence intervals; asterisks indicate statistical significance ($*p < 0.05$, $**p < 0.01$).

While ADOS-2, IQ, and SCQ-A are standardized tests used for a comprehensive diagnosis of autism, the co-occurrence of medical, psychological, and psychiatric conditions constitutes a primary source of phenotypic heterogeneity, modulating the clinical presen-

tation of autism related features throughout the developmental trajectory. Our analysis highlighted a specific divergence in OCD & related symptomatology. Specifically, the negative coefficient for the female group indicates that autistic boys were significantly more likely to present with Obsessive-Compulsive Disorder and related symptomatology compared to girls ($\beta = -1.76$, $SE = 0.81$, $z = -2.18$, $p = 0.029$).

Additionally, a marginal trend was observed for distress symptomatology, which also tended to be more frequent in boys ($\beta = -0.85$, $p = 0.079$), although this is significant only at 10%. Conversely, no significant sex-based differences were identified for other symptomatology, such as anxiety ($\beta = -0.27$, $p = 0.52$), mood/depression ($\beta = 0.05$, $p = 0.93$), or attentional symptomatology ($\beta = -0.75$, $p = 0.10$). These results indicate a relatively shared co-occurrence profile across sexes in our sample, with the notable exception of compulsive behaviors, which appear to be more characteristic of the male phenotype in this cohort.

Taken together, our exploratory analysis reveals a nuanced landscape where broad similarities in total scores ($p > 0.05$) coexist with specific phenotypic differences between autistic girls and boys. While the overall autistic-related behaviors and cognitive performance appear comparable across sexes, the internal architecture differs: the female phenotype in this sample is characterized by preserved eye contact with reduced facial expressivity and specific lower performance in processing speed and visuoperceptual abilities and fluid reasoning tasks. Conversely, the male phenotype is distinguished by a significantly higher prevalence of overt compulsive symptomatology (OCD) ($p < 0.05$) and specific distress patterns. This suggests that sex differences in autism are not global, but rather embedded in specific behavioral-cognitive sub-domains and co-occurring health profiles.

These preliminary group-level differences support the relevance of exploring whether distinct behavioral-cognitive profiles can be identified beyond sex-based comparisons. Accordingly, we next conducted a cluster analysis using the variables that showed significant variation, in order to uncover potential latent subgroups within the sample.

3.4.2 Diagnostic Delay and Topological Predictors

To better understand diagnostic timing in autism, we conducted regression models using clinical, cognitive, and network-derived variables as predictors of later diagnosis (logistic regression) and diagnostic age in years (linear OLS regression). Table 3.2 summarizes the principal results of the linear model of diagnosis age, incorporating key characteristic domains, WISC-V and WAIS-IV Full-Scale IQ and IQ subscales, and interaction terms. Also table B.2 shows the results table of logistic model to the probability of later diagnosis (see B).

Model 2 establishes a baseline, confirming that, on average, girls in our sample were diagnosed approximately 2.5 years later than boys ($p < 0.01$). Model 2 exposes the resultant coefficient when adding clinical control variables. However, the addition of network topology metrics in Model 3 significantly improved the model fit (Adjusted R^2 rises from 0.175 to 0.340), reducing residual standard errors. This improvement reveals that the

structure of features explains variance that characteristics counts alone cannot.

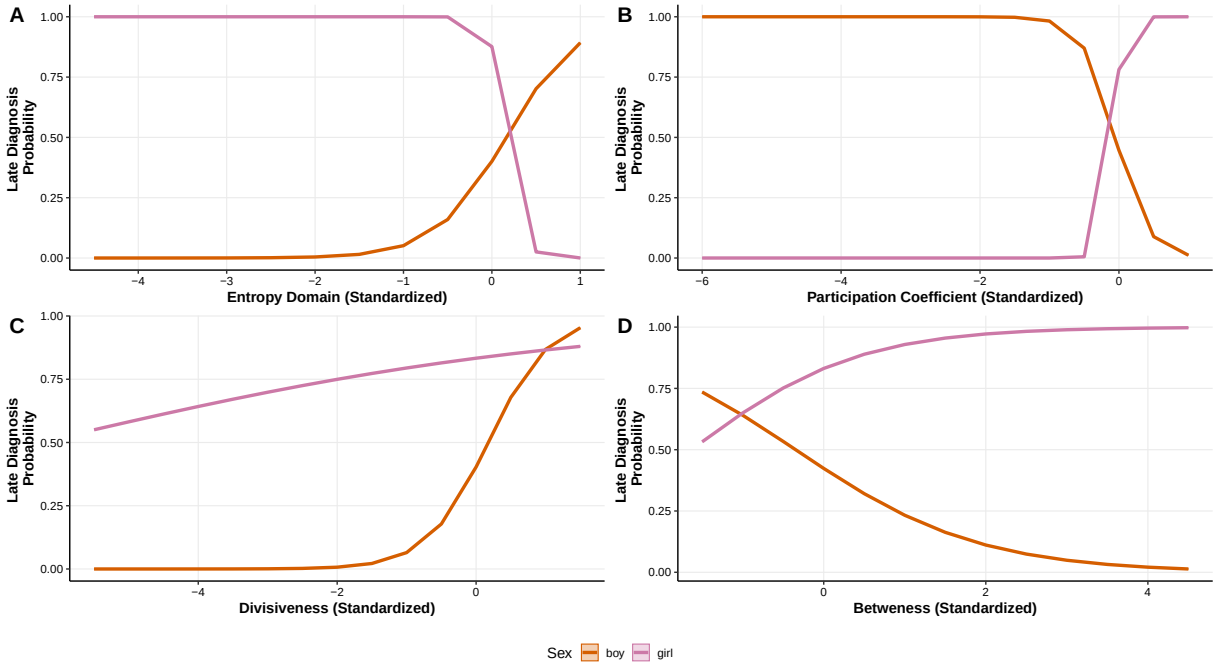

Figure 3.2: **Sex-Specific Interactions between Network Topology and Late Diagnosis Probability. A-D.** Predicted probability of late diagnosis (age > 8 years) as a function of standardized topological metrics, estimated via logistic regression models with sex-interaction terms. Metrics were derived from a bipartite network constructed from the co-occurrence of clinical features (ADOS-2 items) within each individual. **(A)** Domain Entropy: Quantifies the dispersion of features across clinical domains; higher values indicate more diffuse or "noisy" phenotypic profiles. **(B) and (D)** Participation Coefficient and Betweenness: Measures of network connectivity and integration. They reflect the extent to which features are interconnected and the capacity of nodes to act as bridges between distinct phenotypic domains. **(C)** Divisiveness: Measures the degree of fragmentation in the individual's features network.

In other words, girls are less likely to receive a timely diagnosis, even when exhibit similar characteristics to those observed in boys. However, this diagnostic delay is not uniform across all girls. Our analysis sought to identify which factors might reduce or increase this gap. In particular, the logistic regression highlights the role of co-occurring psychological symptomatology, which may blur the clinical picture.

When girls present attentional symptomatology, the probability of later detection decreases. This effect is driven by the interaction term ($\beta = -2.689, p < 0.05$), which effectively reverses the global trend observed in the baseline model ($\beta = 1.655, p < 0.01$). Such symptomatology appear to counterbalance the baseline delay typically seen in girls. Consequently, while attentional symptomatology are associated with a diagnostic delay in the general sample, they act as an early detection signal specifically for girls, resulting in a net negative effect on the age of diagnosis.Conversely, girls without such features remain at greater risk of delayed identification. Furthermore, boys with attentional symptomatology presents a higher age of diagnosis comparative with boys without that symptomatology and with girls.

Regarding restricted and repetitive behaviors (ADOS-2) and general behavioral symptomatology, these factors affect diagnostic timing similarly across both sexes, showing no

	Dependent variable:		
	Diagnosis Age (Years)		
	(1)	(2)	(3)
Girl	1.406*	2.558***	2.452***
	(0.722)	(0.855)	(0.809)
Attentional Symptomatology		1.585*	1.655*
		(0.952)	(0.860)
Repetitive Restringed Behavior (ADOS-2)		−0.832***	−0.988***
		(0.219)	(0.220)
Behavior Symptomatology		−1.061	−1.664**
		(0.725)	(0.714)
Absent Pretend Play (SCQ-A)		1.106	1.462**
		(0.729)	(0.674)
Divisiveness			2.806***
			(0.783)
Domain Entropy			0.770
			(1.892)
Betweenness			−0.731
			(0.445)
Participation Coef.			−3.567*
			(2.012)
Girl * Attentional Sympt.		−1.788	−2.689*
		(1.464)	(1.352)
Girl * Divisiveness			−1.761
			(1.522)
Girl * Domain Entropy			−13.512*
			(7.769)
Girl * Betweenness			1.436*
			(0.808)
Girl * Participation Coef.			17.895*
			(9.367)
Constant	8.398***	9.115***	9.604***
	(0.500)	(0.904)	(0.826)
Observations	96	94	94
R^2	0.039	0.228	0.440
Adjusted R^2	0.029	0.175	0.341
Residual Std. Error	3.536	3.259	2.912
DF	df = 94	df = 87	df = 79
F Statistic	3.790*	4.285***	4.441***
DF	(df = 1; 94)	(df = 6; 87)	(df = 14; 79)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3.2: **Regression of Diagnostic Age.** OLS regression comparing clinical baseline vs. network-augmented models. OLS regression models predicting age of diagnosis (in years). Model (1) establishes a baseline using only sex. Model (2) adds clinical controls. Model (3) integrates bipartite network topology metrics. The inclusion of topological variables significantly improves model fit (Adjusted R2 increases to 0.341). Key findings include the baseline penalty for female sex and the accelerating effect of repetitive behaviors. Regarding topology, divisiveness shows a general positive association with diagnostic age (delay) across both sexes. Crucially, Model (3) reveals sex-specific interactions: structural coherence (high betweenness and participation coefficient) is associated with increased diagnostic delay for girls (acting as a barrier), whereas systemic disorder (high Domain Entropy) and attentional symptomatology significantly mitigate this delay.

interaction with sex. In both groups, externalizing signals contribute to earlier identification, specifically, restricted and repetitive behaviors ($\beta = -0.98, p < 0.01$) and general behavior symptomatology ($\beta = -1.66, p < 0.05$) trigger early ascertainment.

As mentioned before, an innovation, on the regression of diagnosis age as dependent variable, is the aggregation of independent variables that comes from a bipartite network between children and features (for details see chapter 3.3). The table 3.2 demonstrates that when adding topological metrics, such as divisiveness, betweenness, domain entropy, and participation coefficient, the goodness of fit of the model improves; both the R^2 and adjusted R^2_{adj} values improve and residual standard errors also decrease. Specifically, network divisibility serves as a reference structural predictor ($\beta = 2.80, p < 0.01$), indicating that a fragmented network structure generally predicts diagnostic delay across the population, regardless of sex. This reflects a direct relationship between the lack of structural connectivity among symptomatic groups and the timing of formal identification (Figure 3.2), where symptoms remaining isolated in disconnected clusters are associated with a later age of diagnosis for both boys and girls.

In addition, network topology modulates the diagnosis age through a significant crossover interaction between network metrics and sex. The topological configuration can be summarize as two opposing patterns: the connectivity association and the entropy moderator. The metrics domain entropy and participation coefficient presents a significant interaction with sex (at %90 of confidence).

While high network participation (integration of features) generally predicts earlier diagnosis in the baseline ($\beta = -3.56$, $p < 0.1$), the significant positive interaction for girls ($\beta = 17.89$, $p < 0.1$) reverses this effect. This suggests that for girls, having a highly integrated feature network hinders timely identification.

Conversely, the interaction between female sex and domain entropy is significantly negative ($\beta = -13.51$, $p < 0.1$). This indicates that a high-entropy phenotypic profile, where features are scattered across domains rather than cleanly organized, contributes to diagnosis in girls. In essence, girls need a "noisier" or less structured presentation to overcome the diagnostic barrier.

Contrary to the intuitive expectation that a clear, organized profile aids diagnosis, our results indicate the opposite for autistic girls: a coherent architecture (high participation coefficient) obscures the diagnosis, whereas a "messy", high-entropy profile with externalizing aids (attentional symptomatology) is required to render the female autism phenotype visible to the clinical gaze (see figure 3.2).

To visualize which specific features drive the centrality effect described above, we projected the network metrics onto a radial layout. The radial Figure 3.3 aims to display the differentiated betweenness centrality among the feature nodes of the bipartite network. It can be observed that the vast majority of autism-related behavior nodes are located in the center, illustrating their high betweenness centrality. On the other hand, co-occurring medical and psychological conditions are located far away from the center, indicating their lower betweenness centrality. IQ nodes show diverse centrality values; there are many nodes located at the center, but many others are placed at the second ring and a

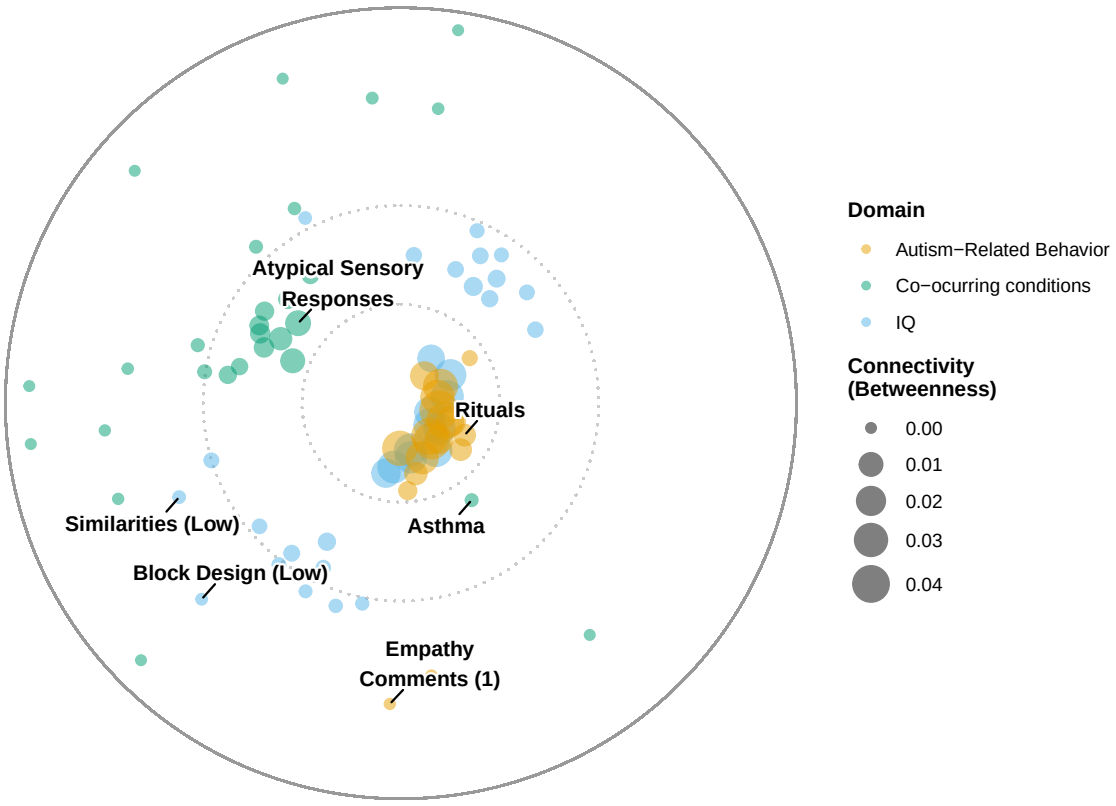

Figure 3.3: **Domain Betweeness Centrality.**. This radial network projection visualizes the structural role of clinical features within the bipartite network. Nodes represent specific clinical characteristics, colored by domain. Node Size is proportional to betweenness centrality, indicating the capacity of a node to bridge different phenotypic domains. The inner circle contains features with high connectivity (e.g., repetitive behaviors). This high integration is associated with an increased probability of late diagnosis in autistic girls. In contrast, the periphery (outer ring) comprises features with lower integration (e.g., Block Design).

few are sited at the periphery.

The interaction with female sex and betweenness centrality predicts a increase in the age of diagnosis. That result indicate that, in autistic girls greater centrality is associated with longer diagnostic timelines.Autistic girls show a critical dependence on a higher burden of autism-related characteristics, co-occurring medical and psychological symptomatology, and lower cognitive performance. The age of diagnosis for girls is strongly positively correlated with participation coefficient (purple line in Figure 3.2 B). Notably, girls with low participation coefficient present a drastically reduced probability of early detection compared to their male counterparts.

Interestingly, while the Hub Score provides a robust measure of overall 'Clinical Load,' it did not emerge as a significant predictor of diagnostic delay. This suggests that the timing of diagnosis is not primarily driven by the cumulative intensity of autism-related characteristics, but rather by the topological arrangement and cross-domain integration of these symptoms.

Together, the results of regression models indicate that, in autistic girls, greater peripheral dispersion is associated with earlier detection, whereas greater core centrality is associated with longer diagnostic timelines. Autistic girls with "central" profiles characterized by average IQ and subtle autism-related social traits faced a scenario where conformity to the core autism-related features was associated with hindered timely diagnostic identification.

Our model predicts divergent trajectories based on this structural configuration. A hypothetical girl presenting with a highly organized symptom structure, where features are tightly interconnected (high participation) and centered (high betweenness), accompanied by average attention skills—is statistically projected to face the most delayed diagnostic.

In contrast, a girl with the same features but a distinct structural arrangement, characterized by "systemic noise" (high entropy), loose connections between domains (low participation), and the presence of attentional symptomatology is predicted to receive a diagnosis significantly earlier.

For autistic boys, the diagnostic trajectory is fundamentally different. They exhibit a "ceiling effect" in probability of detection (as seen in the orange lines of Figure 3.2), where the probability of timely diagnosis remains consistently high regardless of their connectivity or entropy levels. For a boy, having a coherent, central profile is not associated with delayed diagnosis; on the contrary, it often aids identification. The only significant structural risk factor for boys is high divisiveness (fragmentation), which acts as a global barrier, obscuring the diagnostic signs of autism under the shadow of disconnected co-occurrences.

This structural dynamic 3.2 suggests that while boys exhibit a diagnosis that likely regardless of network structure, girls require a specific configuration to overcome the recognition barrier. This suggests that, for boys, diagnosis is associated with a broad range of clinical profiles, including those with lower levels of core centrality.

3.4.3 Multilayer Network Architecture

Our network analysis uncovers fundamentally divergent structural embeddings that go beyond simple count-based differences. A core finding is the differential topological organization of characteristics: while the global network structure remains statistically stable across sexes (NCT: M, $p = 0.96$; global strength S, $p = 0.27$), the internal wiring exhibits a significant meso-scale reorganization. This suggests that the architecture of the interactions between boys and girls differs substantially in how distinct layers communicate.

To quantify these roles, we analyzed Node Versatility, a metric that measures how diverse a node's connections are across different layers. High versatility identifies features that function as multi-domain hubs—connecting, for example, cognitive performance with medical status—whereas low versatility suggests features that operate primarily within their specific domain.

The analysis of node versatility reveals that identical clinical features assume distinct

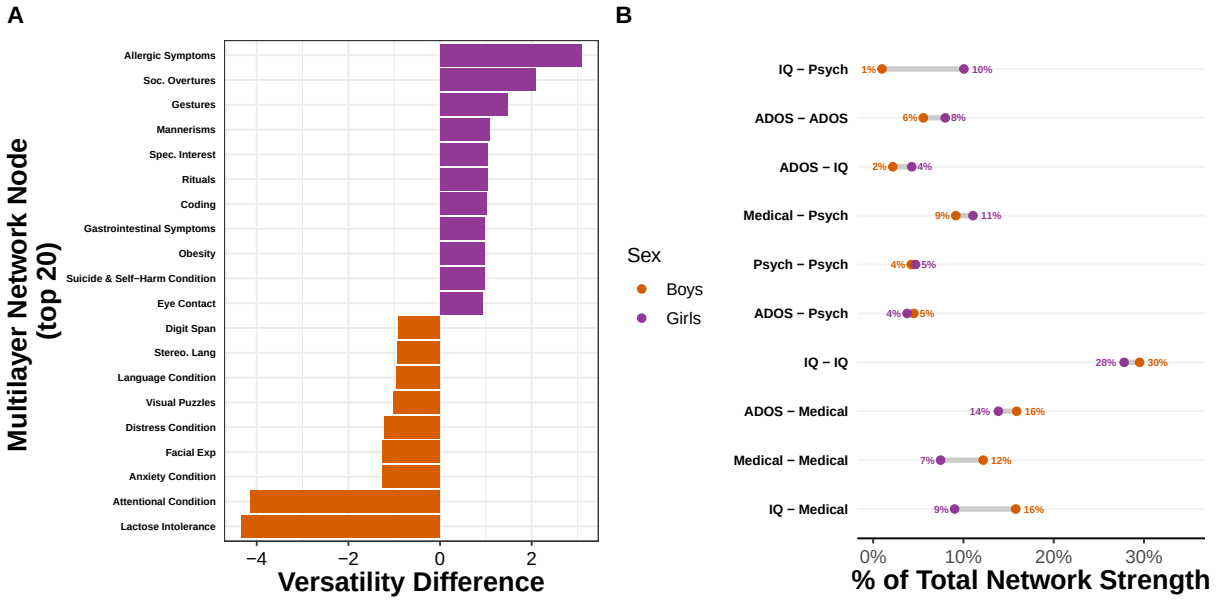

Figure 3.4: **The Multilayer Phenome: Architecture of Symptom Integration. A** Differential node versatility (Sex gap). A diverging bar chart ranking the top 20 features with the largest versatility difference between sexes. Orange bars indicate traits where males exhibit higher network versatility (e.g., attentional conditions, lactose intolerance), whereas purple bars identify traits acting as dominant bridges in the female network (e.g., allergic symptoms, social overtures), confirming a sex-specific prioritization of characteristics integration. **B** Inter-layer connectivity strength. A dumbbell plot comparing the proportion of total network strength allocated to interactions within and between specific layers (IQ, ADOS-2, Medical, Psych) for males (orange) and females (purple). This reveals sex-specific structural dependencies; for instance, females may show distinct connectivity burdens in medical-psychological couplings compared to the male baseline.

functional roles depending on sex. In the female network, specific nodes act as structural bridges that actively couple disparate domains. Specifically, we observed that the ADOS-2 items Reciprocal Communication and Gestures in girls function as connectors linking language-related cognitive capacities with internalizing co-ocurrence, such as obsessive–compulsive and mood symptomatology.

Notably, girls exhibited greater participation and versatility in cognitive nodes—especially Coding, Vocabulary, and Similarities, that integrated tightly with co-occurrences (dense Cyan connectivity). In contrast, boys showed stronger intra-layer connectivity, reflected in lighter node fills. For example, nodes such as Facial expressions and Unusual sensory interest, remained modular, topologically central within the Autiism-related Behavior layer but lacking the cross-domain integration seen in females.

To visualize these topological differences, we mapped the network properties using two key metrics for the visual encoding: node strength and bridging ratio (see B to definition and formula details). The figure 3.5 highlight the principal bridges features (look at B.1 in section B to see the complete network structure). Node size is proportional to total node strength, representing the centrality of the feature within the entire system. Crucially, node fill (grayscale) encodes the bridging ratio (B_i), a metric that quantifies the proportion of a node's strength derived from inter-layer connections. Darker nodes ($B_i \to 1$) act as structural bridges coupling distinct layers, while lighter nodes ($B_i \to 0$) function as locally integrated units within their own layer. Furthermore, the edges are colored to explicitly track cross-domain integration: Cyan links represent coupling between Cognitive Performance and Co-occurring layers, Purple connects Cognitive to Autism-Related Characteristics, and Yellow links Autism-Related Characteristics to Co-occurring conditions.

Visual inspection of Figure 3.5 immediately reveals contrasting topologies. The female network (Fig. 3.5A) is characterized by a dense web of Cyan and Purple inter-layer connections, with numerous dark nodes appearing in both the Cognitive and Co-occurring layers. This indicates a system where cognitive traits and physical symptoms are highly interdependent. In contrast, the male network (Fig. 3.5B) appears more modular, particularly within the Autism-Related Behavior layer, where nodes are lighter (indicating lower bridging). However, the male structure is not disconnected; rather, it shows a specific integration (Yellow and Cyan lines) anchoring physical medical co-ocurring condition to specific cognitive domains, suggesting a different, more "backbone-like" structural organization compared to the diffuse integration seen in females. Specifically, in the male network, the cognitive layer is directly anchored to the co-occurring layer through specific physical health —such as Gastrointestinal issues, Obesity, and Asthma. This targeted coupling suggests that these manifestations are integrated into the male autistic phenotype through a distinct regulatory axis, rather than the widespread, cross-domain interdependence observed in females.

Expanding the analysis to inter-layer connectivity reveals that for autistic girls, the network displays an architecture characterized by a massive reorganization of connectivity towards the psychological axis. Quantitatively, the structural coupling between the Cognitive and Psychological layers was stronger in girls (9% of total strength) compared to boys (2%).

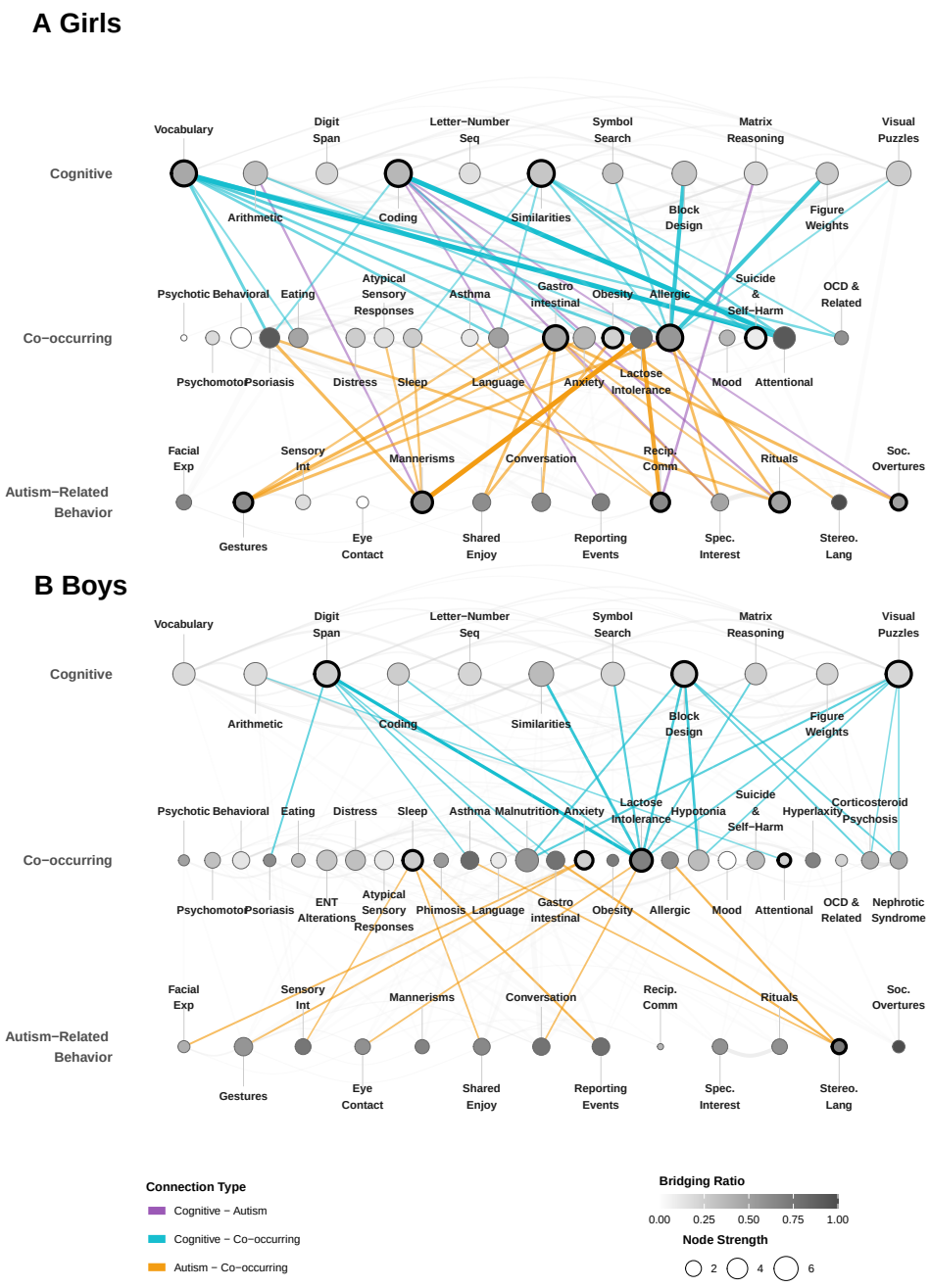

Figure 3.5: **Sex-Differential Multilayer Network Architecture.** The networks visualize the structural coupling between Cognitive (top), Co-occurring (middle), and Autism Core (bottom) layers for (A) Girls and (B) Boys. Edges: Links represent significant Spearman correlations ($|r| > 0.25$). Edge thickness and opacity are proportional to the correlation strength (thicker lines indicate stronger statistical associations). Edge color encodes the type of inter-layer connection (e.g., Cyan links Cognitive nodes to Co-occurring conditions); intra-layer connections are visualized in faint grey to reduce visual noise and highlight cross-domain integration. Nodes: Node Size indicates Total Node Strength (the sum of all connection weights for that node). Node Fill (Grayscale) represents the Bridging Ratio (the proportion of a node's strength that comes from inter-layer connections). Darker nodes (ratio $\rightarrow 1$) function as structural bridges connecting distinct layers, while lighter nodes (ratio$\rightarrow 0$) are locally integrated within their own layer.

Furthermore, the connection between medical and psychological layers was also stronger in girls (12% vs. 9%), indicating a structural coupling where physical conditions are topologically linked to emotional distress (figure 3.4 B). Conversely, the network of boys exhibits a profile where physical symptomatology act as structural correlates of autism related behavior. Boys showed a stronger direct coupling between the autism-related behavior and the medical layers (16% vs. 13%), as well as a more prominent cognitive-medical interface (16% vs. 12%).

Finally, the identification of converging hubs highlights distinct profiles. In autistic girls Suicide & Self-Harm, Gastrointestinal, and Allergic, emerge as high-traffic nodes integrating inputs from cognitive layers. This positions physiological and emotional dysregulation at the center of the female network. In contrast, the male network is anchored by high-versatility hubs related to arousal and regulation, in particular anxiety and attentional symtomatology. Unlike the female pattern in which anxiety is linked to cognitive items Vocabulary and similarities, in boys, these features maintain strong ties to block design, visual puzzles, and coding.

3.4.4 Exploratory Cluster Analysis and Topological Segregation

Based on the previous statistical results, the exploratory cluster analysis included those behavioral and cognitive variables that showed statistically significant sex differences, i..e., four standardized WISC-V subtests [Visual Puzzles (VP), Vocabulary (VOC), Coding (CD), and Symbol Search (SS)], and three ADOS-2 algorithm items (Unusual Eye Contact, Facial Expressions Directed Toward the Examiner, and Stereotyped or Idiosyncratic Use of Words or Phrases). The final PAM model ($k = 4$) identified four distinct behavioral-cognitive profiles, each presenting a unique phenotypic signature.

The optimal partitioning around medoids model revealed that, as illustrated in Figure 3.6, the resulting clusters delineate distinct trajectories of how autistic girls group within the sample and the variables from IQ and ADOS-2 used to generate them:

Cluster 1 ($n = 34$) grouped individuals with a solid cognitive performance across all WISC-V subtests (mean BAL = 11, VOC = 10.2) and a predominance of score 2 (atypical) in Unusual Eye Contact (B1). The group consisted mostly of boys (n = 26), and nearly all participants showed marked atypicality in the B1 item, which aligns with specific social challenges.

Cluster 2, ($n = 19$) was characterized by average-to-high WISC-V performance but a notably high proportion of score 0 in Unusual Eye Contact (all 17 individuals), suggesting no atypical eye contact. This cluster had a majority of girls ($n = 15$).

Cluster 3 , ($n = 22$) exhibited relatively high WISC-V scores and a consistent pattern of score 2 in Unusual Eye Contact. Crucially, all participants received a score of 1 on Item A4, reflecting subtle language-related atypicalities. This group showed a balanced sex distribution (11 boys, 11 girls), representing a mixed phenotype defined by linguistic nuances rather than sex-specific traits.

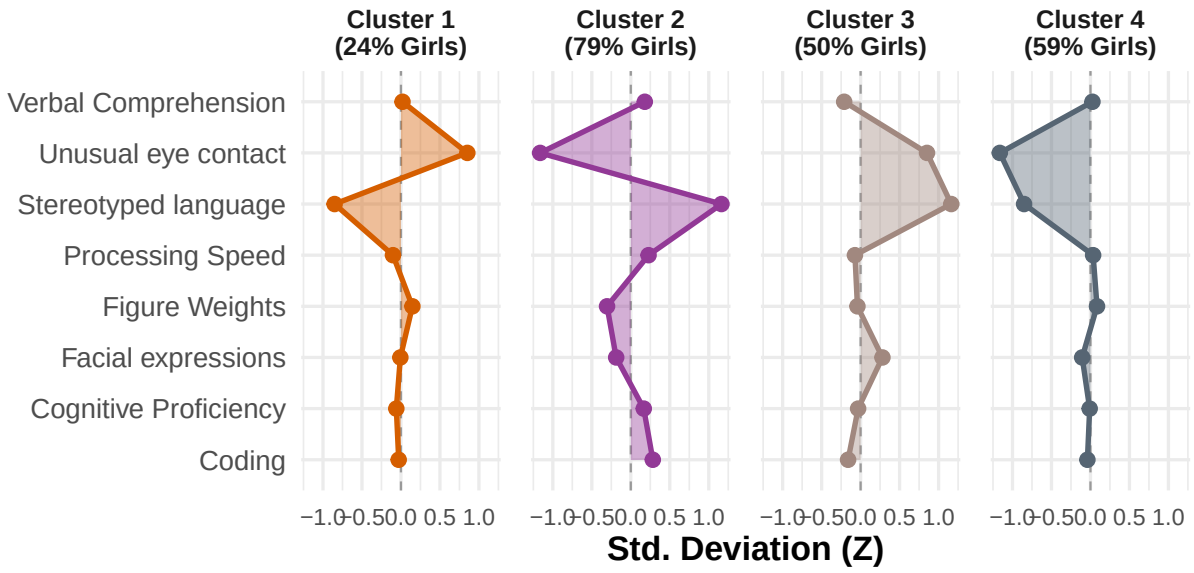

Figure 3.6: **Sex distribution differs significantly across clusters** ($\chi^2 = 16.55, p < .001$). The radar plot illustrates the mean scores for each cluster across the variables. All data are presented as standardized Z-scores, allowing for the comparison of different scales by placing them on a common axis of standard deviations from the study's average. Positive Z-scores indicate that a characteristic is more frequent or prominent in that cluster compared to the overall population, while negative Z-scores reflect a lower prevalence or relative attenuation. **Cluster 1:** It exhibits positive Z-scores in Unusual eye contact and Figure Weights, while showing a pronounced negative Z-score in Stereotyped language. Regarding Verbal Comprehension, Facial expressions, Cognitive Proficiency, and Coding, this cluster aligns closely with the sample average ($Z \approx 0$). **Cluster 2:** It shows the highest positive Z-scores in Verbal Comprehension, Processing Speed, Cognitive Proficiency, and Coding. Simultaneously, it presents negative Z-scores in Unusual eye contact, Figure Weights, and Facial expressions, indicating a lower prevalence relative to the sample mean. **Cluster 3:** It shows a generalized trend toward positive Z-scores in almost all variables, with the exceptions of Coding and Verbal Comprehension, which present negative deviations. **Cluster 4:** It presents negative Z-scores in Unusual eye contact ($Z < -0.5$), Stereotyped language, and Facial expressions. The remaining variables stay near the zero-line.

Cluster 4, ($n = 22$) displayed moderate-to-lower cognitive performance compared to the other three clusters (mean scores ~ 10). While they presented with no atypical eye contact (score 0 in B1), the majority showed mild alterations in facial expressions (B2). This cluster, comprising a mixed but female-leaning demographic (59% girls).

The distribution of sex was significantly associated with the identified cluster structure $\chi^2 = 16.548, df = 3, p - value = 0.0008752$, despite being blind to sex during their construction. Specifically, Cluster 2 exhibited a significantly higher proportion of girls compared to the expected distribution, statistically validating it as a female-preponderant phenotype.

As an exercise to understand the patterns across the clusters we test their statistical differences on sex, total Restricted Repited Behavior, and network topology metrics that were excluded from the algorithm. This allowed us to test whether these dimensions would emerge spontaneously as external validators of the identified groups.

The robustness of these clusters was further evaluated using variables external to the

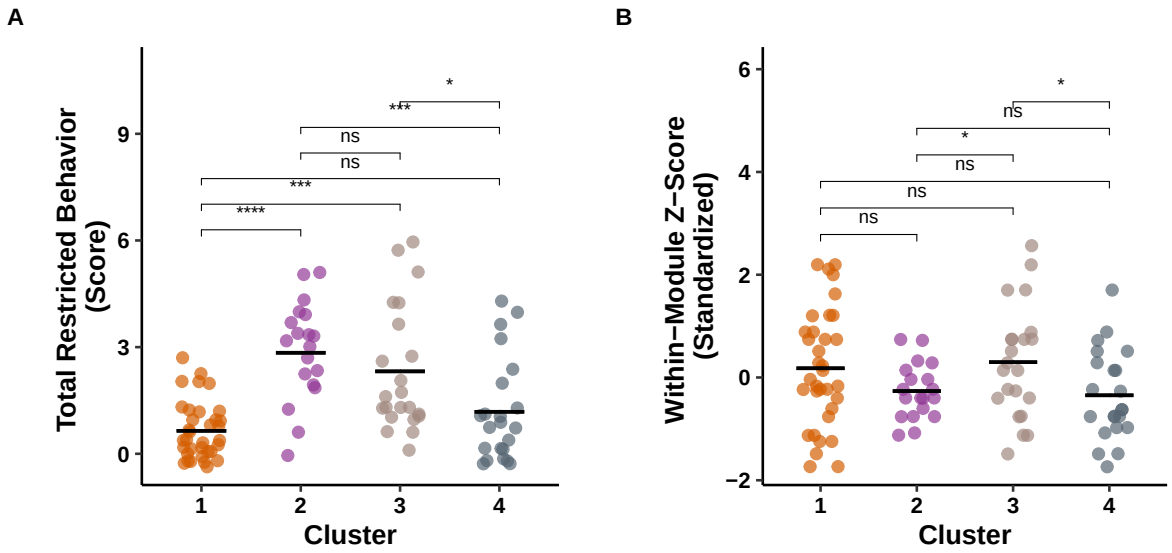

Figure 3.7: **External Validation of clusters.** The strip plots compare the identified clusters across two independent dimensions not used in the clustering algorithm. **(A) Clinical differences:** The Total Restricted Repited Behavior score acts as a strong discriminator between groups (F-test significant, pairwise comparisons as shown). Notably, Cluster 2 exhibits a distinct symptom profile compared to Cluster 1 ($p < .0001$), confirming phenotypical separation based on observable behavior. **(B) Structural Consistency:** In contrast, the Within-Module Z-Score, metric from the bipartite network, reveals structural stability. No significant differences were found between the primary clusters (1 vs. 2, ns), indicating that despite their clinical differences, both represent equally robust and consolidated network phenotypes. Cluster 3 shows the highest cohesion, distinguishing it from the more heterogeneous Cluster 4 ($p < .05$). *Notes:* Black horizontal lines represent the group mean. Significance levels: * $p < .05$, *** $p < .001$, **** $p < .0001$, ns = not significant.

clustering model (Figure 3.7). Clinically, cluster 2 demonstrated significantly lower scores in Total Restricted and Repetitive Behaviors (CRR) compared to Cluster 1 ($p < .0001$; Figure 3.7A). This confirms that the girl phenotype of the sample is a distinct configuration with lower overt repetitive symptomatology.

Structurally, the topological analysis (Figure 3.7B) revealed that despite these clinical differences, Cluster 2 exhibits a Within-Module Z-Score comparable to Cluster 1 (ns). This indicates that the cluster with 79% of girls is structurally as consolidated and robust as the cluster with majority of boys within the symptom network; it is a consolidated core, not a peripheral variation.

However, the highest degree of structural consolidation emerges in Cluster 3. This group displays significantly higher cohesion than both Cluster 2 ($p < .05$) and cluster 4 ($p < .05$). This distinguishes Cluster 3 as a "hyper-specific" linguistic phenotype with tightly interconnected symptoms.

In contrast, the significant difference emerges with Cluster 4, which displays the lowest structural cohesion ($p < .05$ vs Cluster 3). While Cluster 3 appears as a highly specific and cohesive phenotype, Cluster 4 represents a structurally diffuse or peripheral group, likely aggregating individuals with diffuse symptomatology that lacks the topological definition of the other three profiles.

Cluster 4 aligns with the dispersed nodes of the topology. Among girls, diagnostic timing shows distinct associations across cognitive profiles: lower cognitive performance is associated with earlier identification, whereas in profiles with average cognition performance (Cluster 2), earlier identification is observed associated with higher core centrality and greater peripheral dispersion.

While regressions confirm that sex—including its interaction with bipartite topological variables—is a significant predictor of diagnostic age, and multilayer networks explain the distinct configuration and structural integration of features across sexes, cluster analysis identifies specific subgroups that are significantly differentiated by sex based on a selective set of variables. Together, these methods demonstrate that autistic girls present a distinct and structurally coherent phenotypic profile.

3.5 Discussion

We examined sex-specific differences and overlaps in autism-related features in a sample of 97 autistic girls and boys aged 6 to 17 years without intellectual disability, integrating network methodologies with complementary computational approaches to identify possible differentiated profiles. Our findings showed that girls were diagnosed, on average, two years later than boys and exhibited distinct profiles in reciprocal communication, gestures, and specific cognitive indices. By modeling individual–characteristic associations through bipartite and multilayer network frameworks, we identified sex-differentiated global patterns encompassing both clinical characteristics and co-occurring conditions. These results highlight the importance of understanding autism as a complex system of interrelated characteristics rather than as a set of isolated traits, underscoring the need for sex-sensitive diagnostic practices.

A critical contribution of this study is the identification of a structural opposite pattern predicting diagnostic timing. Our bipartite network regression analysis revealed that the internal organization of a child's profile acts differentially according to sex. Specifically, we observed that structural connectivity metrics, namely Betweenness Centrality and Participation Coefficient, exhibit a significant crossover effect. While having a highly interconnected and integrated profile does not delay diagnosis in boys, it acts as a significant barrier for females, extending the diagnostic timeline. In girls, a high Participation Coefficient implies that symptoms are evenly distributed and integrated across domains. We propose that this coherence functions as a structural barrier where the statistical integration of cognitive assets and social subtleties dilutes the autistic signal, rendering the female phenotype less clinically explicit.

Complementarily, our models identified specific factors that facilitate detection. The interaction between sex and Domain Entropy revealed that girls are significantly more likely to receive an early diagnosis when their profile exhibits high systemic disorder. Similarly, the presence of Attentional Symptomatology functioned as a significant facilitator of recognition for females. This implies that for the female phenotype, clinical detection is facilitated when the coherence of the profile is disrupted, either by a high dispersion of symptoms across domains (entropy) or by the presence of salient co-occurring traits like

attention-related conditions. These findings imply that the clinical signal in girls becomes visible primarily when the profile loses its internal integration and presents characteristics of systemic heterogeneity.

In this way, network science contributes not only new analytical tools but also a theoretical lens for understanding the diagnosis of autism as a matter of relational visibility. The empirical utility of this framework is underscored by our results: the inclusion of topological metrics substantially improved model fit (increasing the adjusted R-squared from 0.208 to 0.341). This increase indicates that the topological configuration, or how symptoms are wired together, captures variance in diagnostic timing that the mere presence of individual symptoms cannot explain.

Beyond diagnostic timing, our multilayer network analysis disentangles the internal architecture of these phenotypes. In girls, the regression and clusters reveal that phenotypic centrality acts peculiarly as a trap that delays diagnosis. The multilayer architecture explains this phenomenon by showing a highly integrated network where Reciprocal Communication and Gestures nodes do not function as isolated symptoms but as high-versatility bridges connecting verbal cognitive capacity with the co-occurrence layer, specifically anxiety and mood symptomatology. Unlike boys, where Rituals and Mannerisms are segregated autism-related behaviors, in girls these nodes act as transversal connectors amalgamating behavior with emotional response. This topological configuration creates a profile of internal coherence (Cluster 2) where cognitive and social competence structurally masks the visibility of autism-related traits. Therefore, to be detected early, girls require breaking this harmony through external noise, either through attentional symptomatology or high systemic entropy, as predicted by the regression models.

By contrast, in boys, phenotypic centrality facilitates early diagnosis, a finding supported by the modular architecture of their multilayer network. Here, the autism-related behaviors, such as alterations in Facial Expressions and Sensory Interests, remain segregated within their own layer, generating a clean and easily identifiable clinical signal. Versatility in the male network is not diffuse but concentrated in specific gatekeeper nodes, particularly lactose intolerance, attentional condition, and anxiety symptomatology. These nodes act as anchors connecting the autism-related behavior layer with non-verbal cognitive domains such as visual puzzles and block design. This compartmentalized structure validates the formation of Cluster 1, a profile where the intensity of autism-related behaviors is not diluted by cognitive integration, allowing the standard diagnostic system to recognize this specific male-prevalent phenotype without the need for additional disruptive co-occurring conditions.

The integration of our three computational approaches reveals a coherent narrative regarding the distinct phenotypic pathways. The clustering analysis validated the topological segregation of sex-specific profiles. We identified Cluster 1, comprising mostly boys, characterized by high trait intensity and modularity, which maps onto the rapid-diagnosis trajectory. While girls are distributed across several groups, Cluster 2 (comprising 79% of girls) represents a distinct structural phenotype where male presence is marginal. This group embodies the configuration predicted by our regression models: individuals with high cognitive scores and high structural integration (participation coefficient).This group of girls is similar to those who remain undiagnosed for longer periods, suggesting that their

delayed diagnosis is not due to a lack of characteristics, but rather to the cohesive nature of their symptom architecture. Unlike fragmented or diffuse profiles, such as Cluster 4, this group presents an integrated relational structure that is difficult to diagnose.

Notably, ADOS-2 items such as Reciprocal Communication and Social Overtures did not emerge as significant predictors of diagnostic timing in our regression models. This does not imply that the ADOS-2 lacks validity. The fact that divisiveness predicted timing better than individual ADOS-2 items reinforces that diagnostic delay is driven by the configuration of traits rather than the mere presence of specific social challenges.

Studying these relations directly allows network research to move beyond the question of which features differ between groups and instead investigate how the structure of relationships among characteristics varies across populations. This perspective aligns with the concept of recognitional justice (Fraser, 2000). Diagnostic systems privilege certain structural configurations as recognizable while neglecting others. Network analysis makes this mechanism empirically visible by showing how gendered and cultural patterns of interconnection shape who is identified, when, and under what profile (Diemer et al., 2022).

Our results can contribute to clarifying the features of autistic domains to improve the recognition of autism between sexes by taking into account nuances in nosology, behavioral presentation, developmental change, and contextual constrains (M.-C. Lai & Szatmari, 2020). Differences in presentation and developmental trajectory by sex need to be clarified at the level of narrow constructs or behavioral examples. This type of analysis could contribute to being a first step for future research efforts that contribute not only to diagnostic and therapeutic strategies, but also to the planning of relevant educational strategies in the training of health professionals and public policies in the region.

Chapter 4

General Discussion

This thesis set out to investigate the structural origins of temporal disadvantage, which manifests as lower academic retention and autism diagnostic delay, across two distinct domains. By adopting a Computational Social Science framework, we demonstrated that being late is rarely a random individual accident. Instead, it is often a quantifiable consequence of structural misalignment between individuals and the institutional architectures they navigate.

Integrating findings from higher education and clinical diagnosis, this work advances a unified thesis stating that relational topology acts as a hidden determinant of outcomes. Whether navigating a choice architecture in education or a diagnostic framework in healthcare, the specific configuration of an individual's preferences or features relative to the core structure of the system determines their trajectory.

Central to these findings is a network theory of institutional fit, where success and timely recognition are functions of topological proximity. Classic theories in both education and psychology emphasize fit as a predictor of developmental outcomes (Eccles & Roeser, 2011). In education, Tinto's integration model suggests that persistence depends on the congruence between student motivations and the academic environment. In clinical psychology, accurate diagnosis depends on the fit between the presentation of a patient and nosological prototypes. Our work operationalizes this abstract concept through rigorous topological metrics. In the higher educational retention, we quantified fit as Preference Alignment, showing that students whose preference networks were dispersed, characterized by high network distance, exhibited significantly lower first-year retention rates. This effect persisted regardless of their academic ability, suggesting that informational fit, understood as the coherence of the navigational map of a student, is as crucial as academic readiness for ensuring retention.

Similarly, in the autism domain, we quantified fit through structural connectivity metrics. We identified a topological pattern in females whereby greater phenotypic coherence, marked by high integration between symptoms, was associated with delayed diagnostic timing. Conversely, systemic entropy, or profile disorder, improved the likelihood of recognition. Thus, in higher education, proximity to a coherent cluster of choices supports retention, whereas in autism diagnosis, internal coherence renders girls invisible. Only

when the female profile becomes sufficiently entropic or disordered does it achieve the visibility required for early diagnosis.

Closely related to this topology of fit is the issue of relational visibility, or how institutional systems struggle to process complexity that deviates from standard prototypes. This aligns with network psychometrics literature, which posits that psychiatric or psychological conditions are not latent entities but complex systems of interactions of clinical behaviors and symptoms. In higher education, the admissions system is designed to process vertical sorting based on scores but is blind to horizontal coherence in preference structure. Our Random Forest models revealed that preference distance, a measure of decision-making complexity, was a top predictor of retention, outperforming traditional socioeconomic metrics. This implies the system fails to see the risk inherent in confused or exploratory decision-making until the student fails to be retained.

In autism, the diagnostic frameworks prioritize modular phenotypes. Our multilayer analysis showed that females present a highly integrated architecture in which cognitive and psychological layers are intertwined. Consequently, relational visibility is tuned to the observation of signals that hold intrinsic explanatory value. Systems efficiently process aligned students and modular patients, while those with dispersed preferences (in education) or coherent but subtle clinical networks (in autistic girls) incur a temporal penalty.

These mechanisms of fit and visibility underpin the Silent Architecture of Inequality. While often studied through socioeconomic or biological lenses, our results highlight that inequality is also encoded in relational architectures. In education, we found that the relationship between misalignment and retention interacts with institutional assignment. Access to the first-choice program acts as a buffer that mitigates the risk of misalignment, suggesting that the assignment mechanism itself can either amplify or dampen the risks derived from poor information. The coupling we found between IQ and Psychological Distress in females suggests a structural association within the network. This finding indicates that higher cognitive resources co-occur with increased psychological distress. Further research is needed to better understand this potential trade-off between cognitive functioning and psychological well-being in autistic females.

Methodologically, this dissertation demonstrates the utility of Network Science as a Socioscope, serving as a tool not just for mapping social ties, but for revealing latent phenotypic and decision-making landscapes. Standard variable-based approaches, such as regressions on total scores, would have missed these mechanisms. In Study 1, using average scores hid the fact that preference coherence mattered more than the score itself for many students. In Study 2, comparing mean counts of autism-related behavior failed to explain the diagnostic delay. Only by analyzing the topology, specifically Divisiveness and Centrality, did the sex-specific pattern emerge. This validates the Computational Social Science approach. By treating preferences and features as nodes in a complex system, we can uncover the structural signatures of social problems. Whether it is the fragmented path of a student at risk of non-retention or an undiagnosed girl, the signal is hidden in the network structure.

The cost of being late, comprising the lost years of an undiagnosed autistic girl or the

lost potential of a student leaving higher education, is a product of structural friction. By mapping the relational architectures of education and diagnosis, this work provides the empirical basis for a network-informed policy design. This approach enables identifying at-risk students through preference coherence analysis and complementing diagnostic protocols to recognize the specific network signatures of the female autistic phenotype.

Bibliography

Afzali, M. H., Sunderland, M., Teesson, M., Carragher, N., Mills, K., & Slade, T. (2017). A network approach to the comorbidity between posttraumatic stress disorder and major depressive disorder: The role of overlapping symptoms. *Journal of Affective Disorders*, *208*, 490–496. https://doi.org/10.1016/j.jad.2016.10.037

Aina, C., Baici, E., Casalone, G., & Pastore, F. (2018). The economics of university dropouts and delayed graduation: A survey.

Alchian, A. A. (1950). Uncertainty, evolution, and economic theory. *Journal of political economy*, *58*(3), 211–221.

American Psychiatric Association. (2013a). *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Publishing.

American Psychiatric Association. (2013b). Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc*, *21*(21), 591–643.

Angrist, J. D. (2004). Treatment effect heterogeneity in theory and practice. *The economic journal*, *114*(494), C52–C83.

Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, *113*(27), 7353–7360.

Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic perspectives*, *31*(2), 3–32.

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*(1), 685–725.

Avery, C., & Kane, T. J. (2004). Student perceptions of college opportunities. the boston coach program. In *College choices: The economics of where to go, when to go, and how to pay for it* (pp. 355–394). University of Chicago Press.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Rosenberg, C. R., White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L. C., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., & Dowling, N. F. (2018). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2014. *MMWR Surveillance Summaries*, *67*(6). https://doi.org/10.15585/mmwr.ss6706a1

Barnett, R. (2007). *A will to learn: Being a student in an age of uncertainty*. McGraw-Hill Education.

Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., Young, J.-G., & Petri, G. (2020). Networks beyond pairwise interactions: Structure and dynamics [Defines higher-order structures]. *Physics Reports*, *874*, 1–92.

Begeer, S., Mandell, D., Wijnker-Holmes, B., Venderbosch, S., Rem, D., Stekelenburg, F., & Koot, H. M. (2013). Sex differences in the timing of identification among children and adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(5), 1151–1156. https://doi.org/10.1007/s10803-012-1656-z

Begerson, A. A. (2009). College choice and access to college: Moving policy, research and practice to the 21st century. *ASHE Higher Education Report*, *35*(4), 1–141.

Belasco, A. S., & Trivette, M. J. (2015). Aiming low: Estimating the scope and predictors of postsecondary undermatch. *The Journal of Higher Education*. https://doi.org/10.1353/jhe.2015.0008

Bertanha, M., & Moreira, M. J. (2020). Impossible substitution in fuzzy regression discontinuity designs. *Oxford Bulletin of Economics and Statistics*, *82*(3), 597–629.

Bettinger, E. P., & Baker, R. B. (2014). The effects of student coaching: An evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*, *36*(1), 3–19.

Bischl, B., Richter, J., Bossek, J., Horn, D., Thomas, J., & Lang, M. (2017). mlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions. https://arxiv.org/abs/1703.03373

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Bourdieu, P. (1977). *Reproduction in education, society, and culture*. Sage.

Brannan, D., Biswas-Diener, R., Mohr, C. D., Mortazavi, S., Stein, N., & Brannan, D. (2013). Friends and family: A cross-cultural investigation of social support and subjective well-being among college students. *The Journal of Positive Psychology*, *8*, 65–75.

Brown, C. E., Bernardin, C. J., Beauchamp, M. T., Kanne, S. M., & Nowell, K. P. (2024). More similar than different: Characterizing special interests in autistic boys and girls based on caregiver report. *Autism Research*, *17*(11), 2333–2345. https://doi.org/10.1002/aur.3216

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337–342. https://doi.org/10.1016/j.cognition.2009.05.006

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326.

Candia, C., Encarnação, S., & Pinheiro, F. L. (2019). The higher education space: Connecting degree programs from individuals' choices. *EPJ Data Science*, *8*(1), 39.

Cantor, N., Smith, E. E., French, R. D., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of abnormal psychology*, *89*(2), 181.

Caqueo-Urízar, A., Flores, J., Escobar, C., Urzúa, A., & Irarrázaval, M. (2020). Psychiatric disorders in children and adolescents in a middle-income Latin American country. *BMC Psychiatry*, *20*(1), 104. https://doi.org/10.1186/s12888-020-02512-4

Caruana, N., Brock, J., & Woolgar, A. (2015). A frontotemporoparietal network common to initiating and responding to joint attention bids. *NeuroImage*, *108*, 34–46. https://doi.org/10.1016/j.neuroimage.2014.12.041

Casanova, E. L., Baeza-Velasco, C., Buchanan, C. B., & Casanova, M. F. (2020). The relationship between autism and Ehlers-Danlos syndromes/hypermobility spectrum disorders. *Journal of Personalized Medicine*, *10*(4), 260. https://doi.org/10.3390/jpm10040260

Chade, H., & Smith, L. (2006). Simultaneous search. *Econometrica*, *74*(5), 1293–1307.

Choi, J.-y., & Lee, M.-j. (2023). Score manipulation, density continuity and intent-to-treat effect for regression discontinuity. *Pacific Economic Review*, *28*(4), 552–569.

Chung, W., Jiang, S. F., Paksarian, D., Nikolaidis, A., Castellanos, F. X., Merikangas, K. R., & Milham, M. P. (2019). Trends in the prevalence and incidence of attention-deficit/hyperactivity disorder among adults and children of different racial and ethnic groups. *JAMA Network Open*, *2*(11), e1914344. https://doi.org/10.1001/jamanetworkopen.2019.14344

Cochran, T., & Coles, A. (2012). *Maximizing the college choice process to increase fit & match for underserved students* (Research to Practice Brief) (Retrieved from Institute for Higher Education Policy website). Institute for Higher Education Policy. Washington, DC. https://www.ihep.org/publication/maximizing-the-college-choice-process-to-increase-fit-match-for-underserved-students/

Cook, J., Hull, L., Crane, L., & Mandy, W. (2021). Camouflaging in autism: A systematic review. *Clinical Psychology Review*, *89*, 102080. https://doi.org/10.1016/j.cpr.2021.102080

De Berker, A. O., Rutledge, R. B., Mathys, C., Marshall, L., Cross, G. F., Dolan, R. J., & Bestmann, S. (2016). Computations of uncertainty mediate acute stress responses in humans. *Nature communications*, *7*(1).

Dellapiazza, F., Michelon, C., Rattaz, C., Picot, M.-C., & Baghdadli, A. (2022). Sex-related differences in clinical characteristics of children with ASD without ID: Results from the ELENA cohort. *Frontiers in Psychiatry*, *13*, 998195. https://doi.org/10.3389/fpsyt.2022.998195

Diemer, M. C., & Gerstein, E. (2024). Prior diagnoses and age of diagnosis in children later diagnosed with autism. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-024-06637-3

Diemer, M. C., Gerstein, E. D., & Regester, A. (2022). Autism presentation in female and Black populations: Examining the roles of identity, theory, and systemic inequalities. *Autism*, *26*(8), 1931–1946. https://doi.org/10.1177/13623613221113501

Dworzynski, K., Ronald, A., Bolton, P., & Happé, F. (2012). How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders? *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(8), 788–797. https://doi.org/10.1016/j.jaac.2012.05.018

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of research on adolescence*, *21*(1), 225–241.

Edwards, H., Wright, S., Sargeant, C., Cortese, S., & Wood-Downie, H. (2024). Research review: A systematic review and meta-analysis of sex differences in narrow constructs of restricted and repetitive behaviours and interests in autistic children, adolescents, and adults. *Journal of Child Psychology and Psychiatry*, *65*(1), 4–17. https://doi.org/10.1111/jcpp.13855

Emirbayer, M. (1997). Manifesto for a relational sociology. *American journal of sociology*, *103*(2), 281–317.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Farhat, L. C., Brentani, H., De Toledo, V. H. C., Shephard, E., Mattos, P., Baron-Cohen, S., Thapar, A., Casella, E., & Polanczyk, G. V. (2022). ADHD and autism symptoms in youth: A network analysis. *Journal of Child Psychology and Psychiatry*, *63*(2), 143–151. https://doi.org/10.1111/jcpp.13436

Flum, H., & Blustein, D. L. (2000). Reinvigorating the study of vocational exploration: A framework for research. *Journal of vocational Behavior*, *56*(3), 380–404.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, *65*(6), 591–598. https://doi.org/10.1203/PDR.0b013e31819e7203

for Economic Co-operation, O. O., & Development. (2009). *Education at a glance 2009: Oecd indicators*. ECD Publishing (September 8, 2009).

Fraser, N. (2000). Rethinking recognition. *New Left Review*, *3*, 107–120.

Gaete P, V., & López C, C. (2020). [eating disorders in adolescents. a comprehensive approach]. *Revista Chilena de Pediatría*, *91*(5), 784–793. https://doi.org/10.32641/rchped.vi91i5.1534

Gallegos, J. A., Campos, N. A., Canales, K. A., & González, E. N. (2018). Factores determinantes en la deserción universitaria. caso facultad de ciencias económicas y administrativas de la universidad católica de la santísima concepción (chile). *Formación universitaria*, *11*(3), 11–18.

García, R., Irarrázaval, M., López, I., Riesle, S., Cabezas, M., & Moyano, A. (2021). Encuesta para cuidadores de personas del espectro autista en Chile: Primeras preocupaciones, edad del diagnóstico y características clínicas. *Andes Pediatrica*, *92*(1), 25. https://doi.org/10.32641/andespediatr.v92i1.2307

Gesi, C., Migliarese, G., Torriero, S., Capellazzi, M., Omboni, A. C., Cerveri, G., & Mencacci, C. (2021). Gender differences in misdiagnosis and delayed diagnosis among adults with autism spectrum disorder with no language or intellectual disability. *Brain sciences*, *11*(7), 912.

Ghulam, A., Lei, X., Guo, M., & Bian, C. (2020). Disease-pathway association prediction based on random walks with restart and pagerank. *IEEE Access*, *8*, 72021–72038.

Goldman, S. (2013). Opinion: Sex, gender and the diagnosis of autism—a biosocial view of the male preponderance. *Research in Autism Spectrum Disorders*, *7*(6), 675–679. https://doi.org/10.1016/j.rasd.2013.02.006

González, L. E., & Uribe, D. (2002). Estimaciones sobre la" repitencia" y deserción en la educación superior chilena. consideraciones sobre sus implicaciones. *Calidad en la Educación*, (17), 75–90.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Hamdan-Mansour, A. M., & Dawani, H. A. (2008). Social support and stress among university students in jordan. *International Journal of Mental Health and Addiction*, *6*, 442–450.

Hamilton, A. F. d. C. (2009). Research review: Goals, intentions and mental states: Challenges for theories of autism. *Journal of Child Psychology and Psychiatry*, *50*(8), 881–892. https://doi.org/10.1111/j.1469-7610.2009.02098.x

Hastings, J., Neilson, C. A., & Zimmerman, S. D. (2018). *The effects of earnings disclosure on college enrollment decisions*.

Hiller, R. M., Young, R. L., & Weber, N. (2016). Sex differences in pre-diagnosis concerns for children later diagnosed with autism spectrum disorder. *Autism*, *20*(1), 75–84. https://doi.org/10.1177/1362361314568899

Hlavac, M. (2022). Stargazer: Beautiful latex, html and ascii tables from r statistical output. *The comprehensive R archive network*.

Honneth, A. (1995). *The struggle for recognition: The moral grammar of social conflicts*. Polity Press.

Howe, S. J., Hull, L., Sedgewick, F., Hannon, B., & McMorris, C. A. (2023). Understanding camouflaging and identity in autistic children and adolescents using photo-elicitation. *Research in Autism Spectrum Disorders*, *108*, 102232. https://doi.org/10.1016/j.rasd.2023.102232

Howe, Y. J., O'Rourke, J. A., Yatchmink, Y., Viscidi, E. W., Jones, R. N., & Morrow, E. M. (2015). Female autism phenotypes investigated at different levels of language and developmental abilities. *Journal of Autism and Developmental Disorders*, *45*(11), 3537–3549. https://doi.org/10.1007/s10803-015-2501-y

Hoxby, C. M., & Avery, C. (2012). *The missing" one-offs": The hidden supply of high-achieving, low income students* (tech. rep.). National Bureau of Economic Research.

Hull, L., Mandy, W., Lai, M. C., Baron-Cohen, S., Allison, C., Smith, P., & Petrides, K. V. (2019). Development and validation of the Camouflaging Autistic Traits Questionnaire (CAT-Q). *Journal of Autism and Developmental Disorders*, *49*(3), 819–833. https://doi.org/10.1007/s10803-018-3792-6

Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M. C., & Mandy, W. (2017). "putting on my best normal": Social camouflaging in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *47*(8), 2519–2534. https://doi.org/10.1007/s10803-017-3166-5

Hus, Y., & Segal, O. (2021). Challenges surrounding the diagnosis of autism in children. *Neuropsychiatric Disease and Treatment*, *17*, 3509–3529. https://doi.org/10.2147/NDT.S282569

Isvoranu, A. M., Ziermans, T., Schirmbeck, F., Borsboom, D., Geurts, H. M., & de Haan, L. (2022). Autistic symptoms and social functioning in psychosis: A network approach. *Schizophrenia Bulletin*, *48*(1), 273–282. https://doi.org/10.1093/schbul/sbab084

Jedrzejewska, A., & Dewey, J. (2022). Camouflaging in autistic and non-autistic adolescents in the modern context of social media. *Journal of Autism and Developmental Disorders*, *52*(2), 630–646. https://doi.org/10.1007/s10803-021-04953-6

Jiménez-Molina, Á., Reyes, P., & Rojas, G. (2021). Determinantes socioeconómicos y brechas de género de la sintomatología depresiva en Chile. *Revista Médica de Chile*, *149*(4), 533–542. https://doi.org/10.4067/s0034-98872021000400533

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), 427–431. https://doi.org/10.1038/nature12715

Kana, R. K., Libero, L. E., Hu, C. P., Deshpande, H. D., & Colburn, J. S. (2014). Functional brain networks and white matter underlying theory-of-mind in autism. *Social Cognitive and Affective Neuroscience*, *9*(1), 98–105. https://doi.org/10.1093/scan/nss106

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, *53*(282), 457–481.

Kelly, P. (2001). Youth at risk: Processes of individualisation and responsibilisation in the risk society. *Discourse: studies in the cultural politics of education*, *22*, 23–33.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, *2*(3), 203–271. https://doi.org/10.1093/comnet/cnu016

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, *46*(5), 604–632.

Kochenderfer, M. J. (2015). *Decision making under uncertainty: Theory and application*. MIT press.

Kuh, G. D., Kinzie, J. L., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2006). *What matters to student success: A review of the literature* (Vol. 8).

Lai, M. C., Lombardo, M. V., Ruigrok, A. N. V., Chakrabarti, B., Auyeung, B., Szatmari, P., Happe, F., & Baron - Cohen, S. (2017). Quantifying and exploring camouflaging in men and women with autism. *Autism*, *21*(6), 690–702. https://doi.org/10.1177/1362361316671012

Lai, M.-C., & Szatmari, P. (2020). Sex and gender impacts on the behavioural presentation and recognition of autism. *Current Opinion in Psychiatry*, *33*(2), 117–123. https://doi.org/10.1097/YCO.0000000000000575

Lai, M.-C., Lombardo, M. V., Pasco, G., Ruigrok, A. N., Wheelwright, S. J., Sadek, S. A., Chakrabarti, B., Consortium, M. A., & Baron-Cohen, S. (2011). A behavioral comparison of male and female adults with high functioning autism spectrum conditions. *PloS one*, *6*(6), e20835.

Lambiotte, R., Rosvall, M., & Scholtes, I. (2019). From networks to optimal higher-order models of complex systems [Discusses pathways for flow and higher-order dependencies]. *Nature Physics*, *15*(4), 313–320.

Larroucau, T., Rios, I., Fabre, A., & Neilson, C. (2024). *College application mistakes and the design of information policies at scale*.

Larroucau, T., & Rios, I. (2022). *Dynamic college admissions* (tech. rep.). Tech. rep., Working Paper.(Cited on page 7.)

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., et al. (2009). Computational social science [The foundational manifesto of the field]. *Science*, *323*(5915), 721–723.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355.

Lee, K. S., Gau, S. S.-F., & Tseng, W.-L. (2024). Autistic symptoms, irritability, and executive dysfunctions: Symptom dynamics from multi-network models. *Journal of Autism and Developmental Disorders*, *54*(8), 3078–3093. https://doi.org/10.1007/s10803-023-05981-0

Loomes, R., Hull, L., & Mandy, W. P. L. (2017a). What is the male-to-female ratio in autism spectrum disorder? a systematic review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Loomes, R., Hull, L., & Mandy, W. P. L. (2017b). What is the male-to-female ratio in autism spectrum disorder? a systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*(6), 466–474.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2015). *Autism diagnostic observation schedule, second edition (ADOS-2)*. Western Psychological Services.

Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosta, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S., Esler, A., Grzybowski, A., Hall-Lande, J., & Shaw, K. A. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *MMWR. Surveillance Summaries*, *72*(2), 1–14. https://doi.org/10.15585/mmwr.ss7202a1

Mandell, D. S., Novak, M. M., & Zubritsky, C. D. (2005). Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics*, *116*(6), 1480–1486.

Manski, C. F. (1991). *Adolescent econometricians: How do youth infer the returns to schooling?*

Manski, C. F. (2004). Measuring expectations. *Econometrica*, *72*(5), 1329–1376.

May, K. E., & Kana, R. K. (2020). Frontoparietal network in executive functioning in autism spectrum disorder. *Autism Research*, *13*(10), 1762–1777. https://doi.org/10.1002/aur.2403

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, *142*(2), 698–714.

Merton, R. K. (1968). *Social theory and social structure* (Enlarged). Free Press.

Mitchall, A. M., & Jeager, A. J. (2018). Parental influences on low-income, first-generation students' motivation on the path to college. *The Journal of Higher Education*, *89*(4), 582–609.

Montagut, M., Mas, R. M., Fernández, M. I., & Pastor, G. (2018). Influencia del sesgo de género en el diagnóstico de trastorno de espectro autista: Una revisión. *Escritos de Psicología / Psychological Writings*, *11*(1), 42–54. https://doi.org/10.5231/psy.writ.2018.2804

Moore, R., Owen, L., & Poynton, T. A. (2020). Student preferences for college and career information. *Journal of College Access*, *5*(1), 68–86.

Mundy, P. (2017). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, *47*(6), 497–514. https://doi.org/10.1111/ejn.13720

Muris, P., & Ollendick, T. H. (2021). Selective mutism and its relations to social anxiety disorder and autism spectrum disorder. *Clinical Child and Family Psychology Review*, *24*(2), 294–325. https://doi.org/10.1007/s10567-020-00342-0

Neale, J., & Hudson, L. D. (2020). Anorexia nervosa in adolescents. *British Journal of Hospital Medicine*, *81*(6), 1–8. https://doi.org/10.12968/hmed.2020.0099

Network, N. C. A. (2011). *Maximizing the college choice process to increase fit and match for underserved students*. Institute for Higher Education Policy.

Nieuwoudt, J. E., & Pedler, M. L. (2023). Student retention in higher education: Why students choose to remain at university. *Journal of College Student Retention: Research, Theory & Practice*, *25*(2), 326–349.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism* [Foundational work on encoded categorical biases]. New York University Press.

O'Neill, L. D., Wallstedt, B., Eika, B., & Hartvigsen, J. (2011). Factors associated with dropout in medical education: A literature review. *Medical Education*, *45*, 440–454.

Organización Mundial de la Salud. (2017). *Depresión y otros trastornos mentales comunes*. Organización Panamericana de la Salud.

Pendse, S. R., Lalani, F. M., Munson, S. A., & Sharma, A. (2022). From treatment to punishment: The lived experience of mental health in the sociotechnical systems of care [Discusses how sociotechnical architectures encode bias against divergent profiles]. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–20.

Perez, T., Cromley, J. G., & Kaplan, A. (2014a). The role of identity development, values, and costs in college STEM retention. *Journal of educational psychology*, *106*(1).

Perez, T., Cromley, J. G., & Kaplan, A. (2014b). The role of identity development, values, and costs in college STEM retention. *Journal of educational psychology*, *106*(1).

Posit Team. (2025). *RStudio: Integrated development environment for R* [[Computer software]]. Posit Software, PBC. https://posit.co/

Pucciarelli, F., & Kaplan, A. (2016). Competition and strategy in higher education: Managing complexity and uncertainty. *Business horizons*, *59*(3), 311–320.

R Core Team. (2020). *R: A language and environment for statistical computing* [[Computer software]]. R Foundation for Statistical Computing. https://www.r-project.org

Ratto, A. B., Kenworthy, L., Yerys, B. E., Bascom, J., Wieckowski, A. T., White, S. W., Wallace, G. L., Pugliese, C., Schultz, R. T., Ollendick, T. H., Scarpa, A., Seese, S., Register-Brown, K., Martin, A., & Anthony, L. G. (2018). What about the girls? sex-based differences in autistic traits and adaptive skills. *Journal of Autism and Developmental Disorders*, *48*(5), 1698–1711. https://doi.org/10.1007/s10803-017-3413-9

Rodríguez, M., Rosas, R., & Pizarro, M. (2019, May). *Rendimiento en escala WISC-V en población urbana y rural de Chile* [Manuscript, 1–24].

Roman-Urrestarazu, A., & van Kessel, R. (2022). Inaccurate prevalence estimates impacts autism policy: A letter to the editor in relation to "global prevalence of autism: A systematic review update" by Zeidan et al. (2022). *Autism Research*, *15*(7), 1184–1186. https://doi.org/10.1002/aur.2734

Romero, D. H., Riggs, S. A., & Ruggero, C. (2015). Coping, family social support, and psychological symptoms among student veterans. *Journal of Counseling Psychology*, *62*, 242–252.

Rosas, R., Tenorio, M., Pizarro, M., Cumsille, P., Bosch, A., Arancibia, S., Carmona-Halty, M., Perez-Salas, C. P., Pino, E., Vizcarra, B., & Zapata-Sepulveda, P. (2014). Estandarización de la Escala Wechsler de Inteligencia para Adultos, cuarta edición, en Chile. *Psykhe*, *23*(1), 1–18. https://doi.org/10.7764/psykhe.23.1.529

Rosch, E. H. (1973). Natural categories. *Cognitive psychology*, *4*(3), 328–350.

Russo, F. (2018). The cost of camouflaging autism. *Spectrum News*, 1–9. https://www.spectrumnews.org/features/deep-dive/costs-camouflaging-autism/

Rutter, M., Bailey, A., & Lord, C. (2003). *SCQ. the Social Communication Questionnaire*. Western Psychological Services.

Rynkiewicz, A., Schuller, B., Marchi, E., Piana, S., Camurri, A., Lassalle, A., & Baron-Cohen, S. (2016). An investigation of the 'female camouflage effect'in autism using a computerized ados-2 and a test of sex/gender differences. *Molecular autism*, *7*(1), 10.

Santelices, M. V., Catalán, X., Acevedo, J., Zarhi, M., & Horn, C. (2025). Measuring college information in secondary students and higher education access. *International Journal for Educational and Vocational Guidance*, 1–27.

Santelices, M. V., Zarhi, M., Horn, C., Catalán, X., & Ibáñez, A. (2020). Information sources and transition to higher education: Students, teachers and school counselors perspectives. *International Journal of Educational Research*, *103*, 101617.

Sen, A. (1992). *Inequality reexamined*. Harvard University Press.

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., Lee, L.-C., Rice, C., Giarelli, E., Kirby, R., et al. (2009). Timing of identification among children with an autism spectrum disorder: Findings from a population-based surveillance study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(5), 474–483.

Sipes, M., Matson, J. L., Worley, J. A., & Kozlowski, A. M. (2011). Gender differences in symptoms of autism spectrum disorders in toddlers. *Research in Autism Spectrum Disorders*, *5*(4), 1465–1470. https://doi.org/10.1016/j.rasd.2011.02.007

Smaldino, P. E. (2024). *Modeling social behavior: Mathematical and agent-based models of social dynamics and cultural evolution* [Covers collective search, diversity, and problem-solving constraints (Check if citation refers to the book or a specific 2024 paper like 'Inequality and Collective Search')]. Princeton University Press.

Smith, J., Pender, M., & Howell, J. (2013). The full extent of student-college academic undermatch. *Economics of Education Review*, *32*, 247–261.

Smith, J. V., McQuaid, G. A., Wallace, G. L., Neuhaus, E., Lopez, A., Ratto, A. B., Jack, A., Khuu, A., Webb, S. J., Verbalis, A., Pelphrey, K. A., & Kenworthy, L. (2024). Time is of the essence: Age at autism diagnosis, sex assigned at birth, and psychopathology. *Autism*, *28*(11), 2909–2922. https://doi.org/10.1177/13623613241249878

Solis, A. (2017). Credit access and college enrollment. *Journal of Political Economy*, *125*(2), 562–622.

Sommers, L., Papadopoulos, N., Fuller-Tyszkiewicz, M., Sciberras, E., McGillivray, J., Howlin, P., & Rinehart, N. (2025). The connection between sleep problems and emotional and behavioural difficulties in autistic children: A network analysis. *Journal of Autism and Developmental Disorders*, *55*(4), 1159–1171. https://doi.org/10.1007/s10803-024-06298-2

Soto-Icaza, P., Vargas, L., Aboitiz, F., & Billeke, P. (2019). Beta oscillations precede joint attention and correlate with mentalization in typical development and autism. *Cortex*, *113*, 210–228. https://doi.org/10.1016/j.cortex.2018.12.018

Spiel, K., Haimson, O. L., & Lottridge, D. (2022). Invisibly ill? a critical analysis of disability representation in hci [Critiques normative categories in technological systems]. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–14.

Staiger, D. O., & Stock, J. H. (1994). Instrumental variables regression with weak instruments. *Econometrica*.

Stewart, S., Lim, D. H., & Kim, J. H. (2015). Factors influencing college persistence for first-time students. *Journal of Developmental Education*, *38*, 12–20.

Stiglitz, J. E. (2002). Information and the change in the paradigm in economics. *American Economic Review*, *92*(3), 460–501.

Stock, J. H., & Yogo, M. (2002). Testing for weak instruments in linear iv regression.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Tilly, C. (1998). *Durable inequality*. University of California Press.

Tinto, V. (1993). The dimensions of institutional action. *Leaving College: Rethinking the Causes and Cures of Student Attrition*, 150–160.

To, J. C. S., & Kung, K. T. F. (2025). Sex-typical toy, activity, and playmate preferences in autistic and non-autistic children. *Autism*. https://doi.org/10.1177/13623613251321207

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, *185*(4157), 1124–1131.

Urquiola, M., & Verhoogen, E. (2009). Class-size caps, sorting, and the regression-discontinuity design. *American Economic Review*, *99*(1), 179–215.

Van Heijst, B. F., Deserno, M. K., Rhebergen, D., & Geurts, H. M. (2020). Autism and depression are connected: A report of two complimentary network studies. *Autism*, *24*(3), 680–692. https://doi.org/10.1177/1362361319872373

Vicente, B., Saldivia, S., de la Barra, F., Melipillán, R., Valdivia, M., & Kohn, R. (2012). Salud mental infanto-juvenil en Chile y brechas de atención sanitarias. *Revista Médica de Chile*, *140*(4), 447–457. https://doi.org/10.4067/S0034-98872012000400005

Voelkle, M. C., & Sander, N. (2008). A structural equation approach to discrete-time survival analysis. *Journal of Individual Differences*, *29*, 134–147.

Waldren, L. H., Leung, F. Y. N., Hargitai, L. D., Burgoyne, A. P., Liceralde, V. R. T., Livingston, L. A., & Shah, P. (2024). Unpacking the overlap between autism and ADHD in adults: A multi-method approach. *Cortex*, *173*, 120–137. https://doi.org/10.1016/j.cortex.2023.12.016

Watts, D. J. (2004). The "new" science of networks [Bridging physics-based network science with sociological theory]. *Annual Review of Sociology*, *30*(1), 243–270.

Wood-Downie, H., Wong, B., Kovshoff, H., Mandy, W., Hull, L., & Hadwin, J. A. (2021). Sex/gender differences in camouflaging in children and adolescents with autism. *Journal of Autism and Developmental Disorders*, *51*(4), 1353–1364. https://doi.org/10.1007/s10803-020-04615-z

Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., & Deng, S.-H. (2019). Hyperparameter optimization for machine learning models based on bayesian optimization. *Journal of Electronic Science and Technology*, *17*(1), 26–40.

Ye, X. (2024). Improving college choice in centralized admissions: Experimental evidence on the importance of precise predictions. *Education Finance and Policy*, *19*(2), 308–340.

Zampella, C. J., Parish-Morris, J., Foy, J., Cola, M., Schultz, R. T., & Herrington, J. D. (2024). You should smile more: Population-level sex differences in smiling also exist in autistic people. *Autism*. https://doi.org/10.1177/13623613241301113

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, *15*(5), 778–790. https://doi.org/10.1002/aur.2696

Zeidenberg, M., Jenkins, D., & Calcagno, J. C. (2007). Do student success courses actually help community college students succeed? ccrc brief. number 36. *Community College Research Center, Columbia University*.

Appendix A

Chapter 2 - Supplementary material

A.1 Chilean Supplementary Results

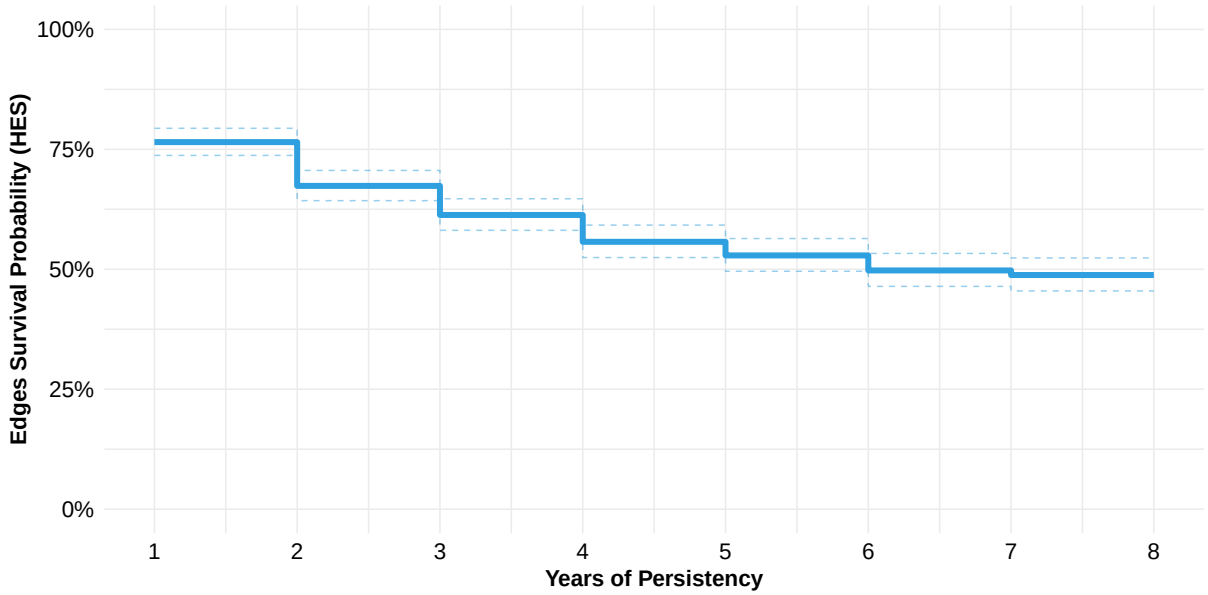

Figure A.1: **Structural Persistence of the Higher Education Space (HES)** This figure displays the Kaplan-Meier survival estimate for the edges connecting generic degree programs. The network was constructed annually following the methodology of Candia et al. (2019), defining edges through the proximity metric (ϕ_{ij}), calculated as the minimum conditional probability of co-application. The x-axis denotes the continuous duration of a link in years since its inception. Dashed lines represent 95% confidence intervals.The plot reveals significant structural inertia; approximately 50% of the connections observed in the base year remain active after 8 years. This indicates a stable "backbone" of complementary career preferences that persists despite short-term market volatility.

	Dependent variable: First-Year Retention		
	Fitted Models		
	Null (1)	Distance (2)	Interaction (3)
Distance		-0.209^{***}	0.178^{***}
		(0.005)	(0.039)
Enrolled Score	0.005^{***}	0.004^{***}	0.005^{***}
	(0.0001)	(0.0001)	(0.0001)
Female	0.096^{***}	0.102^{***}	0.103^{***}
	(0.007)	(0.007)	(0.007)
Age	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
STEM	-0.149^{***}	-0.099^{***}	-0.103^{***}
	(0.013)	(0.013)	(0.013)
High Family Income	-0.064^{***}	-0.066^{***}	-0.068^{***}
	(0.012)	(0.012)	(0.012)
Low Family Income	-0.185^{***}	-0.176^{***}	-0.176^{***}
	(0.010)	(0.010)	(0.010)
Scholarship	0.111^{***}	0.115^{***}	0.115^{***}
	(0.017)	(0.017)	(0.017)
Family High Income * Scholarship	0.025	0.022	0.024
	(0.037)	(0.037)	(0.037)
Family Low Income * Scholarship	0.253^{***}	0.252^{***}	0.252^{***}
	(0.017)	(0.017)	(0.017)
Distance * Enrolled Score			-0.001^{***}
			(0.0001)
Constant	-1.693^{***}	-1.298^{***}	-1.666^{***}
	(0.070)	(0.071)	(0.080)
OECD Area	Yes	Yes	Yes
Year	Yes	Yes	Yes
Father High Education	Yes	Yes	Yes
Mother High Education	Yes	Yes	Yes
School Financial Regime	Yes	Yes	Yes
Province Address	Yes	Yes	Yes
University	Yes	Yes	Yes
Family Size	Yes	Yes	Yes
Type Secondary	Yes	Yes	Yes
Enrolled Preference	Yes	Yes	Yes
McFadden	0.042	0.045	0.045
Cox and Snell (ML)	0.041	0.044	0.044
Nagelkerke (Cragg and Uhler)	0.065	0.069	0.07
Observations	617,622	617,622	617,622
Log Likelihood	$-299{,}975.5$	$-299{,}091.7$	$-299{,}041.9$
Akaike Inf. Crit.	600,207.0	598,441.3	598,343.8
Notes:	*p<0.1; **p<0.05; ***p<0.01		

(continued on next page)

	Dependent variable: First-Year Retention		
		Fitted Models	
	Null (1)	Distance (2)	Interaction (3)

Table A.1: **Logistic regression models predicting first-year retention in the Chilean higher education system.** This table reports the estimated coefficients (log-odds) from three logistic regression specifications. The dependent variable is binary, taking the value of 1 if the student persists in their program for a second year and 0 otherwise. **Model 1 (Null)** includes baseline academic, socioeconomic, and demographic controls. **Model 2 (Distance)** introduces the **Average Preference Distance** variable, revealing a significant negative association with retention ($\beta = -0.209, p < 0.01$), confirming that preference misalignment acts as a risk factor. **Model 3 (Interaction)** adds an interaction term between Distance and Enrolled Score ($\beta = -0.001, p < 0.01$), assessing how academic performance moderates the effect of misalignment. All models control for student characteristics (Gender, Age, STEM enrollment, Family Income, Scholarship status) and include fixed effects for Year, University, School type, and Geographic region (Province/OECD Area). The analysis is based on $N = 617,622$ observations, and network was generated with 5 preferences and a p-value of 0.05. Standard errors are reported in parentheses.

	Dependent variable:					
	First-Year Retention					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	0.203***	0.224***	0.212***	0.200***	0.212***	0.205***
	(0.038)	(0.035)	(0.033)	(0.036)	(0.033)	(0.031)
Enrolled Score	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Female	0.099***	0.099***	0.099***	0.099***	0.098***	0.098***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age	0.001	0.0005	0.001	0.001	0.001	0.0005
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
STEM	−0.114***	−0.116***	−0.118***	−0.115***	−0.117***	−0.118***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
High Family Income	−0.067***	−0.067***	−0.067***	−0.067***	−0.068***	−0.067***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Low Family Income	−0.179***	−0.179***	−0.179***	−0.179***	−0.179***	−0.179***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Scholarship	0.112***	0.113***	0.113***	0.112***	0.112***	0.113***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
High Income * Scholarship	0.024	0.023	0.024	0.024	0.024	0.024
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Low Income * Scholarship	0.253***	0.253***	0.253***	0.253***	0.253***	0.253***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Distance * Enrolled Score	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	−1.712***	−1.732***	−1.732***	−1.711***	−1.730***	−1.731***
	(0.074)	(0.074)	(0.074)	(0.074)	(0.074)	(0.074)
OECD Area	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Father Edu	Yes	Yes	Yes	Yes	Yes	Yes
Mother Edu	Yes	Yes	Yes	Yes	Yes	Yes
Fin. Regime	Yes	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes	Yes
University	Yes	Yes	Yes	Yes	Yes	Yes
Family Size	Yes	Yes	Yes	Yes	Yes	Yes
Type Secondary	Yes	Yes	Yes	Yes	Yes	Yes
Preference	Yes	Yes	Yes	Yes	Yes	Yes
Observations	617,622	617,622	617,622	617,622	617,622	617,622
Network p-value	0.01	0.001	0.0001	0.01	0.001	0.0001
Network Preferences	10	10	10	5	5	5
Log Likelihood	−299,325	−299,369	−299,385	−299,342	−299,399	−299,422
Akaike Inf. Crit.	598,911	598,998	599,031	598,944	599,058	599,103.900
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A.2: **Comparison of Logistic regression models for different networks configurations.** This table reports the estimated coefficients (log-odds) from three logistic regression specifications for first-year retention in the Chilean higher education system. The dependent variable is binary, taking the value of 1 if the student persists in their program for a second year and 0 otherwise. Model 1 to 6 includes the same specification and independent variables as the (3) interaction model specified on Table A.1. The difference between models 1 to 6 regards to the different configuration of the higher education space, variyn the p-value and the amount of preferences from the preference students set takin into account. The results reveals the stability of the specification model and the average preference distance as predictor variable.

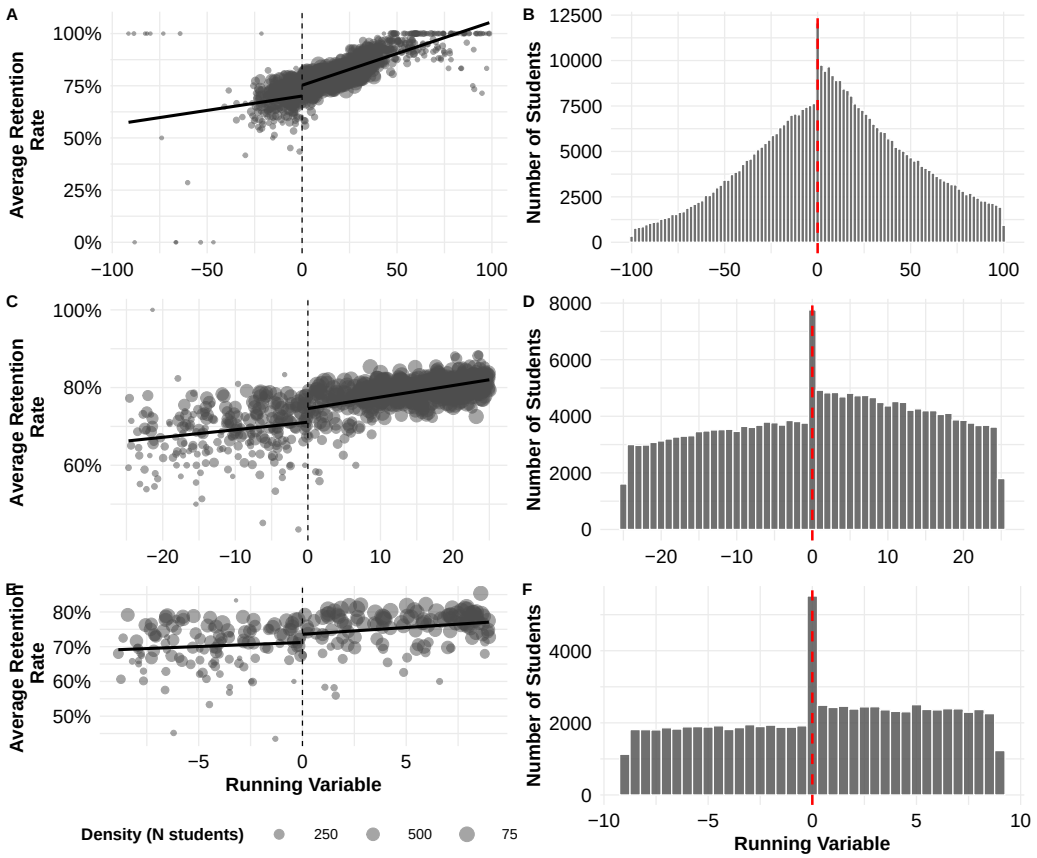

Figure A.2: **Impact of First-Preference Admission on Retention and Design Validity.** This figure displays the graphical analysis of the Regression Discontinuity Design (RDD) across three distinct bandwidth specifications. **(A, C, E) Left Column:** Plots the relationship between the normalized admission score (running variable) and the first-year retention rate. Grey markers represent binned local averages scaled by sample size. Solid lines indicate linear fits estimated separately on each side of the cutoff ($c = 0$), illustrating the causal jump in retention probability for applicants admitted to their first-preference program. **(B, D, F) Right Column:** Displays histograms of the running variable's density. The smoothness of the distribution around the threshold suggests the absence of precise manipulation or sorting by students around the admission cutoff. Rows correspond to decreasing bandwidths to assess local robustness: $h = 100$ (Panels A and B), $h = 25$ (Panels C and D), and $h = 9$ (Panels E and F).

The validity of our causal estimates is fundamentally supported by the large-scale nature of our dataset ($n > 400,000$). Following the asymptotic framework (Calonico et al., 2014), the treatment-effect estimator is consistent provided that the bandwidth satisfies the conditions $h_n \to 0$ and $nh_n \to \infty$. In this study, our sample size allows us to satisfy these conditions effectively; the bandwidth is sufficiently narrow to mitigate asymptotic bias ($B_{n,p}$) while the data density remains high enough to ensure a robust inference with minimal variance ($V_{n,p}$).

The choice of $h = 100$ serves as the structural boundary for the LATE, defining the local subpopulation of interest while maintaining the statistical power required for estimating complex interaction terms. To validate this specification, we contrast our results with the bias-corrected MSE-optimal selector ($h_{CCT} = 8.95$), which is designed to optimize the trade-off between bias and variance. The remarkable convergence between our main specification and the CCT estimator confirms that our results are not an artifact of the window selection but represent a stable causal relationship. This consistency demonstrates that $h = 100$ is localized enough to identify the effect for compliers near the threshold without incurring significant bias across the score distribution.

Method *Criterion*	**CCT** *MSE-Robust*	**IK** *Classic MSE*	**CV** *Cross-Val*	**ROT** *Rule-of-Thumb*	**Manual** *Main Spec*
Bandwidth (h)	8.95	1.40	100	4.50	100
LATE Estimate *(Std. Error)*	$\mathbf{0.222}^{**}$ (0.108)	-0.567^{***} (0.215)	$\mathbf{0.215}^{***}$ (0.041)	0.080 (0.194)	$\mathbf{0.215}^{***}$ (0.041)
Observations	79,422	15,853	447,173	41,760	447,173

Notes: Significance: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A.3: **Sensitivity Analysis of the Local Average Treatment Effect (LATE) to Bandwidth Selection**. This table compares estimates of the Causal Effect (LATE) from the Fuzzy RDD Second Stage using different bandwidth selection algorithms. Standard errors are Cluster-Robust.The robustness analysis confirms the stability of the Causal Effect. Crucially, the CCT optimal bandwidth selector ($h = 8.9$), which minimizes the robust Mean Squared Error, yields a LATE coefficient ($\beta = 0.222$, $p < 0.05$) that is remarkably similar in magnitude to our main specification ($\beta = 0.215$). This convergence between the bias-corrected optimal estimator and our larger bandwidth specification reinforces the validity of the estimated positive causal effect. While the extremely narrow IK bandwidth ($h = 1.4$) shows divergence, likely due to local noise in the immediate vicinity of the cutoff, the consistency across the CV and CCT methods supports the main findings.

To evaluate the sensitivity of our estimates, we replicated the analysis using the MSE-optimal robust bandwidth ($h \approx 9$) proposed by Calonico et al. (2014). The causal effect of treatment remains positive and highly significant ($z = 7.15, p < 0.001$), confirming that the LATE is not an artifact of the bandwidth selection. While the reduced sample size in the optimal window limits the statistical power to identify higher-order interactions, the core coefficients for treatment and preference distance remain consistent in sign and significance with our main specification.

Fuzzy RD Design	**CCT** ($h \approx 9$)	**Int.** ($h = 25$)	**Int.** ($h = 50$)	**Main** ($h = 100$)
Panel A: Second Stage (Outcome: First-Year Retention)				
Treatment Est. (LATE)	1.262^{***}	0.564^{***}	0.383^{***}	0.308^{***}
	(0.080)	(0.044)	(0.033)	(0.030)
Distance	-0.212^{***}	-0.242^{***}	-0.248^{***}	-0.240^{***}
	(0.029)	(0.010)	(0.012)	(0.009)
Running Variable	-0.092^{***}	-0.009^{***}	-0.002^{***}	-0.0003
	(0.010)	(0.002)	(0.001)	(0.0002)
Enrolled Score	0.143^{***}	0.141^{***}	0.116^{***}	0.099^{***}
	(0.031)	(0.023)	(0.015)	(0.013)
LATE Interactions				
Treatment × Distance	0.019	0.060^{***}	0.080^{***}	0.074^{***}
	(0.037)	(0.013)	(0.010)	(0.007)
Treatment × Score	0.201^{***}	0.167^{***}	0.183^{***}	0.196^{***}
	(0.031)	(0.023)	(0.015)	(0.018)
Distance × Score	-0.056^{***}	-0.057^{***}	-0.040^{***}	-0.024^{***}
	(0.020)	(0.011)	(0.005)	(0.005)
Treatment × $D \times S$	0.016	0.025^{**}	0.005	-0.012^{***}
	(0.019)	(0.012)	(0.005)	(0.004)
Panel B: First Stage (Outcome: Enrollment in 1st Pref)				
Z Instrument	1.064^{***}	1.010^{***}	0.888^{***}	0.807^{***}
	(0.221)	(0.295)	(0.244)	(0.234)
Running Variable	0.231^{***}	0.110^{***}	0.069^{***}	0.044^{***}
	(0.007)	(0.003)	(0.002)	(0.002)
Distance	-0.044^{*}	-0.016	0.001	-0.008
	(0.018)	(0.017)	(0.015)	(0.012)
Score	-0.250^{***}	-0.251^{***}	-0.206^{***}	-0.153^{***}
	(0.028)	(0.031)	(0.028)	(0.026)
Distance × Score	0.033^{***}	0.028^{**}	0.026^{**}	0.016^{***}
	(0.010)	(0.010)	(0.009)	(0.007)
Model Diagnostics				
Effective First-Stage F	23.25	11.72	13.24	11.89
Observations	80,274	$\approx 150,000$	$\approx 280,000$	447,173

Table A.4: **Sensitivity of Causal Estimates and Interactions across Bandwidths.**

Note: Cluster-Robust Bootstrap Standard Errors in parentheses. In Panel B, the interaction Distance × Score is included for robustness checks ($h < 100$). Significance: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

A.2 Portugal Supplementary Results

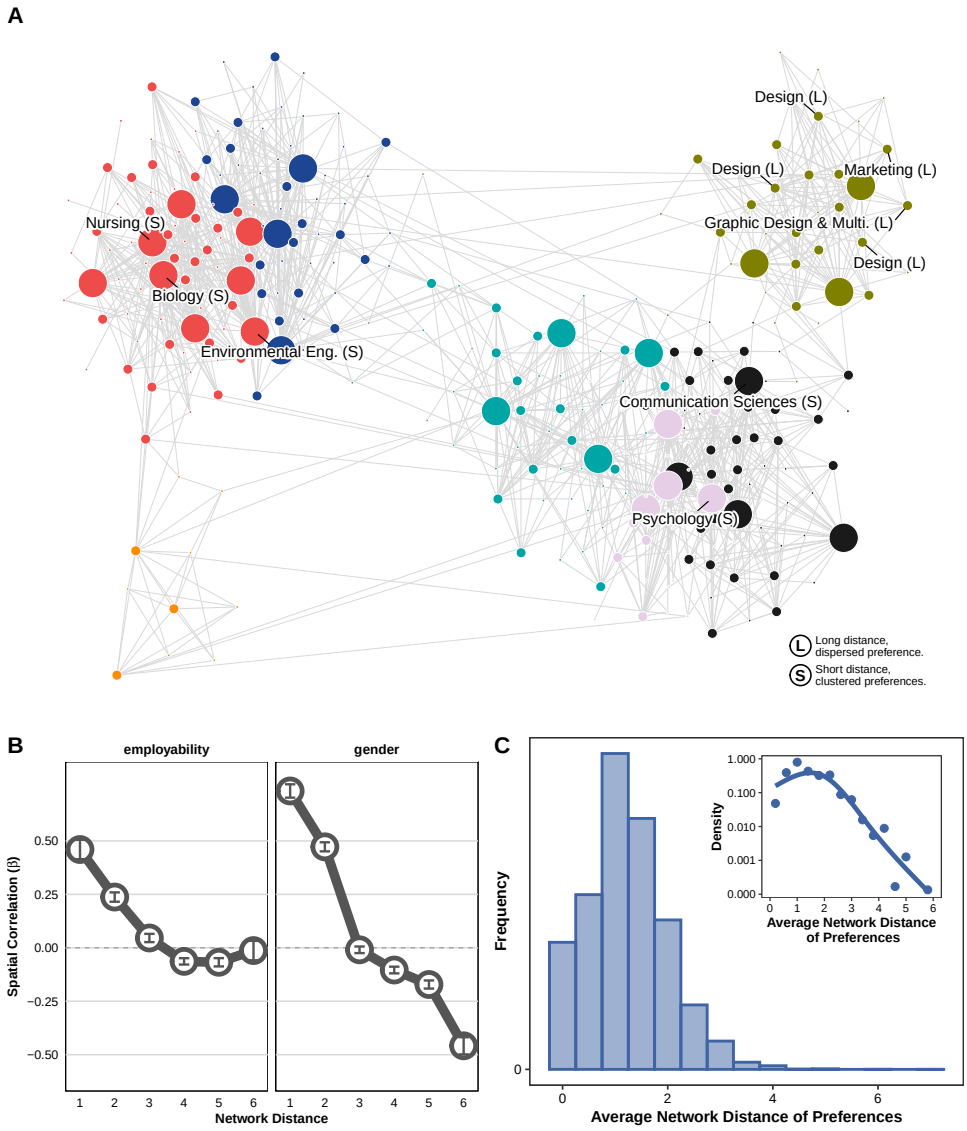

Figure A.3: **Portugal Network structure of degree program preferences and distribution of preference distances.** **A**, Network of degree programs constructed from significant co-occurrence of programs in applicants' ranked preference lists. Nodes represent degree programs, and edges indicate statistically significant co-selection across applicants. Node size reflects the number of applicants listing the program as their first choice. Two regimes of preference alignment are illustrated: clustered preferences with low average network distance (L) and dispersed preferences with high average distance (H). Examples include Design, Graphic design, and Marketing, which occupy distinct network regions. **B**, Spatial correlation between program-level attributes and network distance. Each panel shows how a distinct attribute–gender composition, and employability. Points represent mean correlations across programs and years at each network distance, with error bars indicating standard error. All attributes exhibit a systematic decay as distance increases, suggesting that closely related programs tend to align more strongly in their social and economic characteristics. **C**, Distribution of the Average Preference Distance across applicantsThe inset shows the same distribution on a logarithmic scale, revealing a pattern consistent with exponential decay.

	Dependent variable:		
	First-Year Retention		
	(1)	(2)	(3)
Distance		-0.217^{***}	0.399^{***}
		(0.013)	(0.095)
Enrolled Score	0.003^{***}	0.001	0.006^{***}
	(0.001)	(0.001)	(0.001)
Male	-0.029	-0.037^{*}	-0.036^{*}
	(0.021)	(0.021)	(0.021)
Age	0.073^{***}	0.075^{***}	0.073^{***}
	(0.011)	(0.011)	(0.011)
Enrolled Preference	-0.461^{***}	-0.468^{***}	-0.470^{***}
	(0.006)	(0.006)	(0.006)
Interaction Distance - Enrolled Score			-0.004^{***}
			(0.001)
Constant	1.576^{**}	2.057^{***}	1.295^{*}
	(0.760)	(0.763)	(0.768)
Year	Yes	Yes	Yes
Secondary Program	Yes	Yes	Yes
City	Yes	Yes	Yes
University	Yes	Yes	Yes
Number Of Preferences	Yes	Yes	Yes
McFadden	0.096	0.098	0.099
Cox and Snell (ML)	0.051	0.052	0.052
Nagelkerke (Cragg and Uhler)	0.121	0.124	0.125
Observations	164,716	164,716	164,716
Log Likelihood	$-40{,}549.76$	$-40{,}421.77$	$-40{,}403.13$
Akaike Inf. Crit.	81,501.52	81,247.53	81,212.26
Note:			*p<0.1; **p<0.05; ***p<0.01

Table A.5: **Logistic regression models predicting first-year retention in the Portugal higher education system.** This table reports the estimated coefficients (log-odds) from three logistic regression specifications. The dependent variable is binary, taking the value of 1 if the student persists in their program for a second year and 0 otherwise. **Model 1 (Null)** includes baseline academic, socioeconomic, and demographic controls. **Model 2 (Distance)** introduces the **Average Preference Distance** variable, revealing a significant negative association with retention ($\beta = -0.209, p < 0.01$), confirming that preference misalignment acts as a risk factor. **Model 3 (Interaction)** adds an interaction term between Distance and Enrolled Score ($\beta = -0.001, p < 0.01$), assessing how academic performance moderates the effect of misalignment. All models control for student characteristics (Gender, Age, STEM enrollment, Family Income, Scholarship status) and include fixed effects for Year, University, School type, and Geographic region (Province/OECD Area). The analysis is based on $N = 617,622$ observations. Standard errors are reported in parentheses.

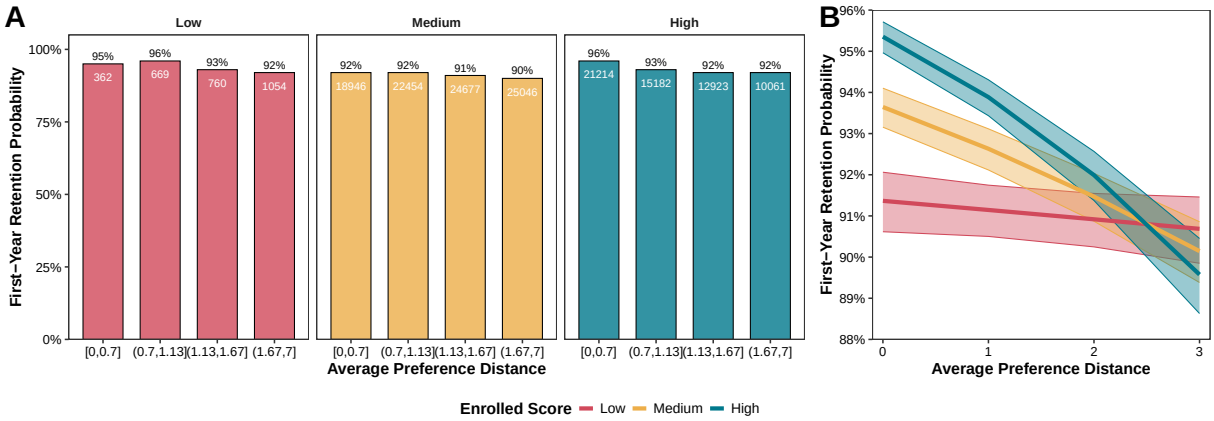

Figure A.4: **A. Observed retention by preference alignment and student quality.** First-year retention probabilities are reported across bins of Average Preference Distance, stratified by tertiles of Enrolled Score. Retention declines monotonically with distance from preferred programs, particularly among students with lower scores. Each bar reports the retention rate and the total number of observations.**B. Logit Predicted probability of first-year retention by preference alignment and enrolled score.** The predicted probability of retention as a function of the average network distance between the degree program students enrolled in and those they initially preferred, stratified by enrolled score (Low, Medium, High), including covariates, shaded bands represent 95% confidence intervals.

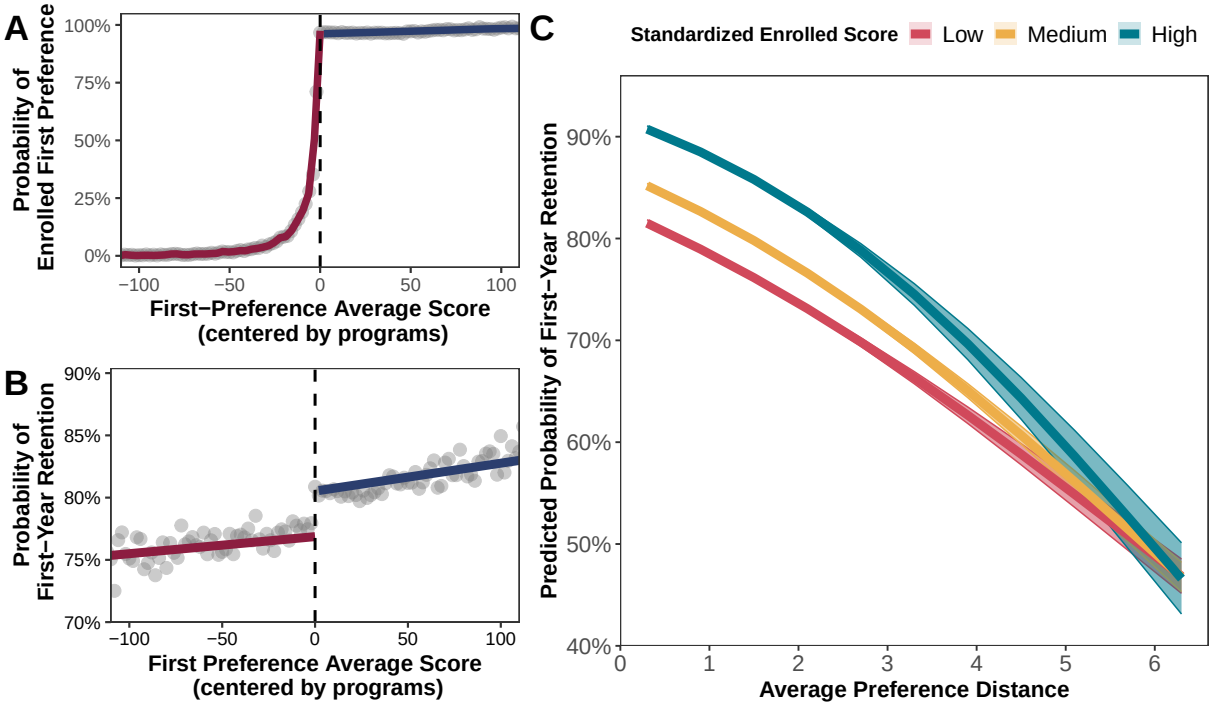

Figure A.5: **Fuzzy regression discontinuity design.A–B. Empirical validation of the fuzzy regression discontinuity design. A.** The probability of enrolling in the first-choice program increases discretely at the cutoff, confirming the presence of a first-stage jump and validating the design as fuzzy rather than sharp. **B.** Retention also increases sharply at the threshold, indicating a causal effect of enrolling in one's top choice on first-year persistence. Together, these panels support the identification assumptions underlying the fuzzy RDD: continuity in potential outcomes and a nonzero first-stage discontinuity in treatment assignment. **C.** The predicted probability of first-year retention by preference alignment and enrolled score (Low, Medium, High), estimates effects using a fuzzy regression discontinuity design; including covariates. Shaded bands represent 95% confidence intervals.

Method *Criterion*	**CCT** *MSE-Robust*	**IK** *Classic MSE*	**CV** *Cross-Val*	**ROT** *Rule-of-Thumb*	**Manual** *Main Spec*
Bandwidth (h)	1.70	1.50	95.0	1.70	100.0
LATE Estimate *(Std. Error)*	**1.702***** (0.168)	**1.711***** (0.164)	**2.002***** (0.106)	**1.702***** (0.168)	**2.002***** (0.106)
Observations	14,812	12,592	128,157	14,812	128,157

Notes: This table compares LATE estimates (Log-Odds) from the Fuzzy RDD Second Stage using different bandwidth selection algorithms for the Portuguese dataset. Standard errors are Cluster-Robust. **Interpretation:** The robustness analysis demonstrates exceptional stability in the estimated Causal Effect. All optimal bandwidth selection methods (CCT, IK, ROT), which suggest a narrow local window ($h \approx 1.5 - 1.7$), yield a positive and highly significant LATE ($\beta \approx 1.70$, $p < 0.001$). Expanding the window to the manual specification ($h = 100$) maintains the direction and significance of the effect ($\beta \approx 2.00$), confirming that the positive impact of first-choice enrollment on retention is consistent across local and global specifications. Significance: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A.6: **Sensitivity Analysis of the Local Average Treatment Effect (LATE) to Bandwidth Selection (Portugal)**

The estimated F-statistic for the first stage is $F \approx 786.35$, which is significantly higher than the most stringent Stock-Yogo critical value of 16.38 for a 10% maximal size distortion. This allows us to reject the null hypothesis of weak instruments, ensuring that our second-stage LATE estimates are robust to identification issues and do not suffer from the asymptotic biases associated with weak instruments (Staiger & Stock, 1994; Stock & Yogo, 2002).

Portugal Fuzzy Regression Discontinuity Design		
Second stage		
Dependent Variable:	First Year Retention (Logit)	
	Coefficient (β)	**Corrected SE**
Treatment Est. (LATE)	$\mathbf{2.062}^{***}$	(0.000)
Distance* (D)	$\mathbf{-0.162}^{***}$	(0.000)
Running Variable (R)	-0.021	(0.000)
Enrolled Score (S)	$\mathbf{0.138}^{***}$	(0.000)
Distance × Enrolled Score	$\mathbf{-0.098}^{***}$	(0.005)
Constant	1.703^{***}	(0.000)
LATE Interactions ($\hat{X}$)		
Treatment × Distance	$\mathbf{-0.097}^{***}$	(0.000)
Treatment × Running	0.014^{***}	(0.000)
Treatment × Enrolled Score	0.219^{***}	(0.000)
First stage:		
Dependent Variable:	Enrollment in First Preference (Probit)	
Running	$\mathbf{0.781}^{***}$	(0.000)
Z Instrument	$\mathbf{1.941}^{***}$	(0.000)
Distance*	0.09	(0.000)
Enrolled Score	$\mathbf{-0.394}^{***}$	(0.000)
Constant	0.049^{***}	(0.000)
Model Diagnostics		
First-Stage F-statistic	1202.4	
Stock-Yogo Critical Value (10% bias)	16.38	
Observations: 447,173 (RDD Window [-50, 50])		

Note: Distance* and Score use standardized variables (std).
Second Stage standard errors are based on Bootstrap inference (Cluster-Robust).
First Stage standard errors are Cluster-Robust (vcovCL).
Significance: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A.7: **Quasi-experimental design the case of Chile** Instrumental Variable Estimation using a Fuzzy Regression Discontinuity Design (FRDD). This table reports the definitive results from the Two-Stage Least Squares (2SLS) estimation strategy of a instrumental variable estimation using a Fuzzy Regression Discontinuity Design (FRDD). Estimated Bootstrap-Corrected Standard Errors, to identify the Causal Effect of enrolling in one's top-ranked program on first-year retention. The first step (Probit/Logit) estimates treatment assignment as a function of the running variable (score relative to program cutoff), the instrument (instrument), Average Preference Distance, and Enrolled Score. The instrument strongly predicts treatment assignment ($\beta = 0.807, p < 0.001$), validating the discontinuity as a source of exogenous variation. The second step (Logit) regresses first-year retention on the predicted treatment status ($\hat{X}$) and covariates. Receiving treatment increases retention by a significant amount ($\beta = 0.308, p < 0.001$), confirming a positive Causal Effect (LATE). The model confirms that Distance (Mismatched Preferences) is a significant predictor of ris ($\beta = -0.240, p < 0.001$). Crucially, the positive coefficient for the Treatment × Distance interaction ($\beta = 0.074, p < 0.001$) suggests that the causal benefit of treatment increases as misalignment (risk) increases, operating as a mechanism of causal mitigation. All reported standard errors are Cluster-Robust and Bootstrap-Corrected, and the model includes multiple interaction terms to control for heterogeneous treatment effects.

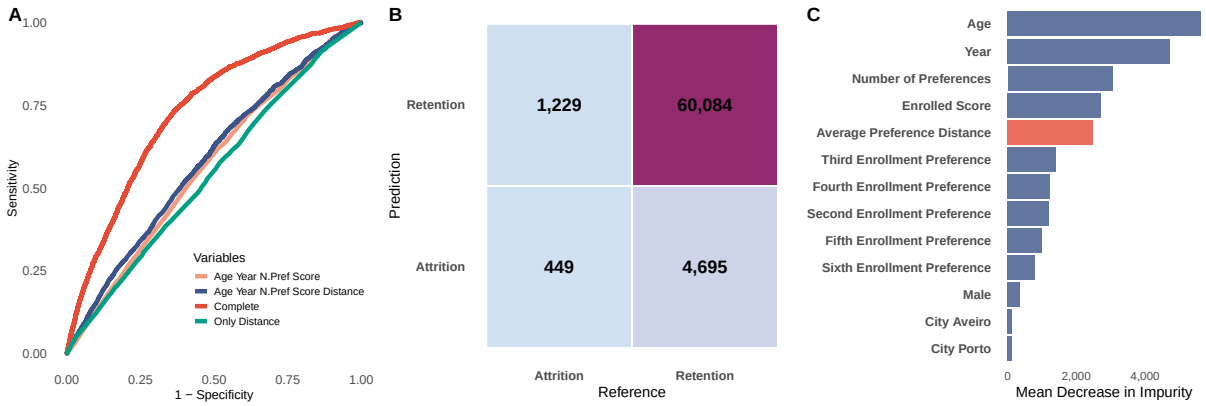

Figure A.6: **Predictive modeling of first-year retention and feature importance. A** Receiver operating characteristic (ROC) curves comparing random forest models for predicting first-year retention using different subsets of predictors. The full model (Complete) achieves the highest discriminative performance, followed by models including Enrolled Preference (a proxy for degree program fixed effects), Age at enrollment, Score Enrolled, and Average Preference Distance. **B** Confusion matrix for the complete model reveals strong predictive accuracy for retention, but lower sensitivity for attrition detection, underscoring the asymmetry of prediction performance across outcomes. **C** Feature importance based on mean decrease in impurity shows that Age, Year, Enrolled Preference, and the newly introduced Average Preference Network Distance are the top five predictors of first-year retention—outperforming traditional academic and socioeconomic variables. Distance is highlighted in orange to underscore its substantive contribution.

Appendix B

Chapter 3 - Supplementary material

Metric (R Variable)	Formula / Definition	Clinical Interpretation
Node Degree (`degree`)	$k_i = \sum_j a_{ij}$	Total count of active symptoms and conditions. A proxy for overall *Symptom Load*.
Symptom Rarity (`rarity`)	$R_i = \sum_j a_{ij} \cdot (1/f_j)$	Weighted sum where symptoms are inversely weighted by their frequency (f_j). High values indicate a child with highly atypical or rare traits.
Divisiveness (`divisiveness`)	$D_i = 1 - \frac{\max(w_{ij})}{\sum w_{ij}}$	Measures profile fragmentation. High values indicate a heterogeneous presentation dispersed across multiple clusters, lacking a dominant focus.
Weighted Entropy (`entropy_w`)	$H_i = -\sum p_j \log_2(p_j)$	Quantifies the uncertainty in the symptom distribution. High entropy suggests a "noisy" or disorganized clinical presentation.
Fiedler Vector (`fiedler`)	Eigenvector of the 2nd smallest eigenvalue of the Laplacian.	Positions each child along a continuous global gradient of the network, capturing the primary topological dimension of variation.
Coreness (`coreness`)	k-core decomposition.	Identifies if a node belongs to the densely connected nucleus of the network. High coreness suggests a "classic" or central clinical presentation.
Hub Score (HITS) (`hits_hub`)	Principal eigenvector of AA^T.	A measure of "Weighted Clinical Load". Identifies children connected to the most "authoritative" (severe/central) symptoms in the network.

Domain Entropy (`domain_entropy`)	Entropy over the distribution of edges across Cognitive, ADOS, and Medical domains.	Measures phenotypic diffusion. High values indicate a profile that spans all three domains equally (diffuse), rather than being concentrated in one.
Participation Coeff. (`participation`)	$P_i = 1 - \sum(k_{i,dom}/k_i)^2$	Quantifies the diversity of connections across domains. High values indicate a multi-systemic profile integrated across cognitive and medical layers.
Betweenness (`bi_betweenness`)	$C_B(v) = \sum \frac{\sigma_{st}(v)}{\sigma_{st}}$	Measures the extent to which a child acts as a "bridge" connecting otherwise disparate symptom clusters.
Harmonic Centrality (`bi_closeness`)	$H(x) = \sum \frac{1}{d(x,y)}$	A robust measure of closeness for disconnected graphs. Indicates how "accessible" the entire symptom space is to the child.
Bipartite Clustering (`bi_clustering`)	Local density of cycles of length 4 (C4).	Reflects local cohesiveness. High values imply the child's symptoms tend to co-occur together in other children (a typical symptom cluster).
Prototypicality (`prototypicality`)	Dot product of child's vector and global symptom popularity.	Indicates how much the child's profile aligns with the "average" or most common presentation in the sample.
Within-Module Z (`within_module_z`)	$z_i = \frac{k_i - \mu_{mod}}{\sigma_{mod}}$	Measures how well-connected a child is specifically within their assigned phenotypic module/cluster.
Projection Strength (`strength_proj`)	Weighted degree in the one-mode projection.	Indicates the intensity of connections to other children based on shared symptoms.

Table B.1: Definitions of topological metrics calculated in the analysis.

	Dependent variable: Later Diagnosis (> 8 years)		
	(1)	(2)	(3)
Girl	0.534	1.530**	1.928**
	(0.416)	(0.615)	(0.875)
Attentional Symptomatology		1.160*	1.397*
		(0.662)	(0.749)
Repited and Restriced Behavior		−0.446***	−0.679***
		(0.157)	(0.210)
Behavior Symptomatology		−0.571	−1.485**
		(0.508)	(0.720)
Absent Pretend Play (SCQ-A)		1.067**	1.680**
		(0.525)	(0.659)
Divisiveness (Standarized)			2.277***
			(0.783)
Domain Entropy(Standarized)			2.517
			(1.809)
Betweenness (Standarized)			−0.886**
			(0.446)
Participation Coef (Std)			−4.232**
			(2.030)
Girl * Attentional Sympt.		−1.519	−2.771**
		(1.021)	(1.268)
Girl * Divisiveness			−2.022
			(1.445)
Girl * Domain Entropy			−13.732*
			(7.725)
Girl * Betweenness_std			1.864*
			(1.120)
Girl * Participation Coef. (Std)			17.312*
			(9.494)
Constant	−0.000	0.028	0.700
	(0.283)	(0.625)	(0.784)
Observations	96	94	94
Log Likelihood	−64.959	−55.166	−43.761
Akaike Inf. Crit.	133.917	124.331	117.522
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.2: **Regression models for late diagnosis age.** Binomial Logit generalized regression models predicting later diagnosis (>8 years). Model (2) establishes a baseline using only clinical and demographic predictors. Model (3) integrates bipartite network topology metrics. The inclusion of topological variables significantly improves model fit (reduction of AKAIKE). Key findings include the baseline penalty for female sex ($\approx$ 2 years delay), the accelerating effect of repetitive behaviors, and the emergence of coreness as a general predictor of diagnostic delay ($p < 0.01$). Crucially, model (2) reveals the sex-specific interactions, in which centrality (page-rank metric) is associated with increased diagnostic delay for girls, while divisiveness mitigates it.

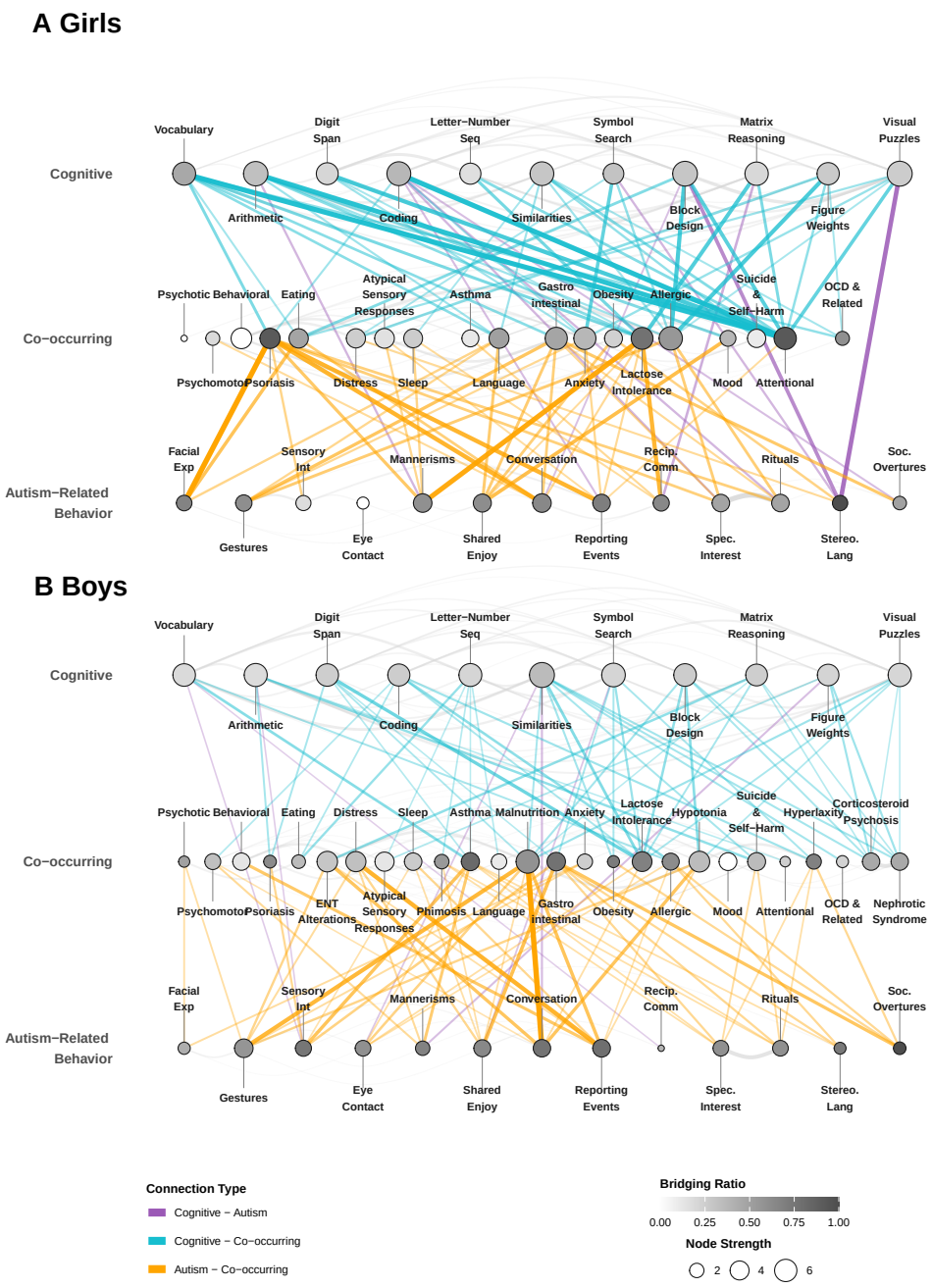

Figure B.1: **Sex-Differential Multilayer Network Architecture.** The networks visualize the structural coupling between Cognitive (top), Co-occurring (middle), and Autism Core (bottom) layers for (A) Girls and (B) Boys. Edges: Links represent significant Spearman correlations ($|r| > 0.25$). Edge thickness and opacity are proportional to the correlation strength (thicker lines indicate stronger statistical associations). Edge color encodes the type of inter-layer connection (e.g., Cyan links Cognitive nodes to Co-occurring conditions); intra-layer connections are visualized in faint grey to reduce visual noise and highlight cross-domain integration. Nodes: Node Size indicates Total Node Strength (the sum of all connection weights for that node). Node Fill (Grayscale) represents the Bridging Ratio (the proportion of a node's strength that comes from inter-layer connections). Darker nodes (ratio $\rightarrow 1$) function as structural bridges connecting distinct layers, while lighter nodes (ratio$\rightarrow 0$) are locally integrated within their own layer.

Metric	Formula / Definition	Clinical Interpretation in Multilayer Context
Node Versatility	$V_i = \sum_{\alpha \neq \beta} w_i^{\alpha\beta}$	Defined as the sum of weights of edges connecting a node in one layer (e.g., Cognition) to nodes in different layers (e.g., Medical). High versatility identifies "bridge nodes" that facilitate cross-domain integration.
Participation Coefficient	$P_i = 1 - \sum_{L=1}^{3} \left(\frac{s_{i,L}}{S_i} \right)^2$	Measures the diversity of a node's connections across the three layers (Cognitive, Autism Core, Co-occurring). A value close to 1 indicates the symptom interacts equally with all domains; a value close to 0 indicates it is isolated within its own layer.
Bridging Ratio	$B_i = \frac{S_{inter}}{S_{total}}$	The proportion of a node's total strength derived from inter-layer connections. Used in visualization to distinguish "locally integrated" nodes (light color) from "structural bridges" (dark color).
Global Inter-layer Density	$D_{inter} = \frac{\sum \lvert E_{inter} \rvert}{\sum \lvert E_{total} \rvert}$	The proportion of the total network energy (strength) allocated to connections between different layers. Higher density indicates a less modular, more systemically integrated phenotype.
Inter-layer Coupling Strength	$S_{L1-L2} = \sum w_{ij} \mid i \in L1, j \in L2$	The total weight of connections between two specific layers (e.g., IQ vs. Anxiety). Used to quantify the specific structural dependency between domains.

Table B.3: **[Multilayer Network Metrics.**Definitions of topological metrics utilized specifically in the Multilayer Network Analysis.