Modeling Strategic Risk in School Choice: A Case for Transparent Design

Mayesha Tasnim[1], Paul Verhagen[1], Tobias Blanke[2], Erman Acar[1,2], Sennay Ghebreab[1]

[1]Institute of Informatics, University of Amsterdam
[2]Institute for Logic, Language, and Computation, University of Amsterdam
{m.tasnim, p.verhagen, t.blanke, e.acar, s.ghebreab}@uva.nl

Abstract

School choice algorithms, which assign students to schools based on preferences, are often designed to balance matching efficiency with strategyproofness. In Amsterdam, policymakers introduced a more efficient assignment mechanism in 2019 that permitted preference manipulation. As strategic behavior increased, transparency around the algorithm was reduced—motivated by prior studies showing that additional information about matching methods tends to increase manipulation. Yet transparency is also essential for institutional trust and informed decision-making. We test whether transparency about the consequences of strategic behavior deters or encourages manipulation. We conduct a behavioral experiment embedded in a survey of 140 parents participating in Amsterdam's 2025 school choice process. Parents selected among truthful, mildly strategic, and aggressively strategic rankings across three scenarios with varying degrees of probability disclosure. While transparency increased the rate of strategic behavior, truthful reporting remained the most common choice under all conditions. To explain these patterns, we estimate a discrete-choice model grounded in Security-Potential/Aspiration (SP/A) theory, which models how parents weigh aspirational goals against the risk of poor outcomes. The model fits observed behavior better than expected utility and prospect theory benchmarks, even though individual covariates are not statistically significant. We find that parents whose children applied to competitive school tracks (e.g., pre-university) were more risk-averse, selecting safer strategies even when aggressive options offered high chances of success. These results challenge the assumption that transparency inherently worsens inequality, and suggest that communicating risk may serve as a viable policy lever for fairness in algorithmic systems.

Introduction

School choice mechanisms co-determine access to educational opportunity and, in turn, long-term social mobility (Hastings and Weinstein 2008). In school choice, assignment algorithms often operate as one-sided matching mechanisms, where students rank schools but schools do not rank students (Abdulkadiroğlu, Pathak, and Roth 2009). Policymakers and researchers have long debated the role of transparency in these systems: while making the rules of the system visible can improve trust and accountability, it is also feared to invite "gaming" that undermines fairness. This tension is especially acute in matching algorithms that are not *strategyproof*, where families can potentially benefit from misreporting their preferred schools (Abdulkadiroğlu, Pathak, and Roth 2009). For example, strategyproof algorithms like Random Serial Dictatorship (RSD) encourage truth-telling, but lead to inefficient outcomes, leaving many students without access to top-ranked schools (Abdulkadiroğlu and Sönmez 1998). In contrast, more efficient matching methods are manipulable and incentivize families to misrepresent their preferences (Bogomolnaia and Moulin 2001; Featherstone 2020; Aksoy et al. 2013).

This balance between efficiency and equity has shaped school assignment policies in many cities. The prevailing assumption is that parents with a high socioeconomic status (SES) are more likely to identify and exploit strategic advantages, while lower-SES families may struggle to do so (Abdulkadiroğlu et al. 2006; Hastings and Weinstein 2008; Hastings, Van Weelden, and Weinstein 2007). To avoid reinforcing inequality, many systems retain strategyproof mechanisms despite reduced efficiency, or limit transparency when using non-strategyproof mechanism. Despite the existence of this assumption, little is known in the literature about how families actually respond to incentives under manipulable systems. Previous work has shown that strategic behavior increases when mechanisms are manipulable, particularly when assignment data or algorithmic rules are made transparent (Pais and Pintér 2008; Calsamiglia, Haeringer, and Klijn 2010). However, these findings are largely derived from experimental studies in lab settings and may not capture how real-world parents behave, particularly in a high-stakes context like school choice, where parents can likely be more risk-averse (He 2012).

Amsterdam offers a relevant case. In 2016, the city adopted a lottery-based mechanism that prioritized strategyproofness but led to dissatisfaction due to inefficiency, with many students matched to schools ranked low on their preference list (Tasnim et al. 2024; Khaddari 2022). Following complaints and legal appeals, a 2019 policy change introduced a placement guarantee, which also rendered the system manipulable. The change was not widely communicated, and strategic knowledge on its manipulability spread informally. To mitigate this, the school boards reduced trans-

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

parency by no longer publishing anonymized preference data or the lowest lottery number at which a student secured a place at each school, which had previously helped parents estimate admission probabilities (Hulshof 2025).

This shift also reflects a common policy assumption derived from the literature: that reducing information transparency deters manipulation by increasing uncertainty. Yet opacity has its own drawbacks, including reduced accountability and trust in systems (Zerilli, Bhatt, and Weller 2022). Moreover, in many matching settings, it is difficult to determine an optimal strategy, even when providing high levels of information transparency (Troyan 2022; Tasnim et al. 2024). This raises a key question: rather than incentivizing manipulation, can transparency about the probabilities of success of strategies serve as a deterrent by making the risks of manipulation more explicit?

To investigate this, we conduct a behavioral experiment embedded in an online survey of 141 parents whose children are about to participate in Amsterdam's lottery-based school choice system. This population was chosen intentionally as real-world stakes may shape how parents perceive and respond to strategic risk, in contrast to typical lab settings, with perceived manipulability being different than actual manipulability. Participants were asked to complete a simplified school matching task under a manipulable algorithm, choosing between being truthful, mildly strategic, and aggressive manipulation. They repeated the task under three sequential conditions: (1) no probability information, (2) full information with high risk, and (3) full information with reduced risk. This experimental setup is designed to allow us to isolate the effects of transparency and strategic risk.

Our findings show that transparency does increase the rate of strategic behavior, but that most parents opt for mild deviations. Aggressive manipulation tactics are not adopted universally, even when risk is low. Truthful reporting remains the most commonly picked strategy in all three scenarios, and among those who deviate, mild strategies are far more common than aggressive ones. To explain this, we model parents' decisions using the Security-Potential/Aspiration (SP/A) framework from behavioral decision theory. In this model, each strategy is evaluated based on its security (avoiding bad outcomes) and potential (achieving one's aspiration), moderated by a parent-specific risk preference. We estimate this model using a multinomial logit and recover interpretable heterogeneity in behavior. The fitted risk preferences are bimodally distributed, and show that parents with children in a more competitive school type tend to be risk-averse, which runs counter to the common assumption that competition fuels strategic behavior.

Overall, this paper makes three contributions. First, we provide empirical evidence on how parents respond to strategic incentives and transparency under a manipulable mechanism. Second, we introduce SP/A theory as a behavioral framework for understanding strategy use in school choice, showing that aspiration-driven behavior aligns with observed patterns. Third, we offer a policy insight: transparency, rather than uniformly enabling manipulation, may reduce harmful strategies by clarifying the risks under manipulable matching methods. Our findings suggest that sharing accurate information about strategy success and failure could discourage harmful manipulation without sacrificing institutional trust. This aligns with recent simulation-based work showing that only a small fraction of strategic reports lead to better outcomes in Amsterdam, implying that explicit transparency could act as a behavioral deterrent (Tasnim et al. 2024). Our findings suggest that transparency can support both equity and efficiency in the design of school choice algorithms.

Related Work

Algorithmic Transparency

With the growing prevalence of algorithmic school allocation, there has been increased scrutiny of how such systems align with public values and requirements that come with the use of algorithms in the public sector.

Transparency in particular is emphasized in key regulatory frameworks such as the EU AI Act (Parliament and Council 2023) and the Digital Services Act (Parliament and Council 2022). The latter, for example, mandates in-scope service providers to publish an annual transparency report with respect to content moderation, and disclose algorithmic operations that may impact the public domain. However, while transparency is increasingly seen as a normative requirement for public sector algorithms, some scholars have argued that great algorithmic transparency carries the risk of strategic manipulation and gaming (Wang et al. 2023). Moreover, there can be legal constraints both from a privacy requirement and an intellectual property requirement that prevent transparency. Crucially for this study, policy choices can also be important variables in determining the level of transparency, including preventing negative media exposure, equity concerns, or fear of strategic manipulation.

This latter argument is central in the Amsterdam school allocation context, with the choice of Random Serial Dictatorship (RSD), orienting around strategic manipulation and the associated fairness concerns. Particularly, this is based on the assumption that greater transparency regarding admission probabilities might incentivize parents to engage in strategic manipulation, exacerbating social inequalities, as parents with a higher socioeconomic background would be more likely to engage in such manipulation (Abdulkadiroğlu et al. 2006; Hastings, Van Weelden, and Weinstein 2007; Hastings and Weinstein 2008). This positions school allocation as not merely a technical problem, but a politically embedded problem that involves political and normative tradeoffs, including between efficiency and fairness, and between transparency and manipulability. At the same time, algorithmic transparency has been shown to have positive effect; prior works suggest it can strengthen trust in algorithmic decision making (Peeters 2020), support actionable recourse (Ustun, Spangher, and Liu 2019) and even improving predictive power of algorithms(Wang et al. 2023). Our study examines whether, within the context of school allocation in Amsterdam, greater transparency on algorithmic outcomes leads to greater (and potentially undesirable) strategic behavior, and whether these effects vary based on parent's

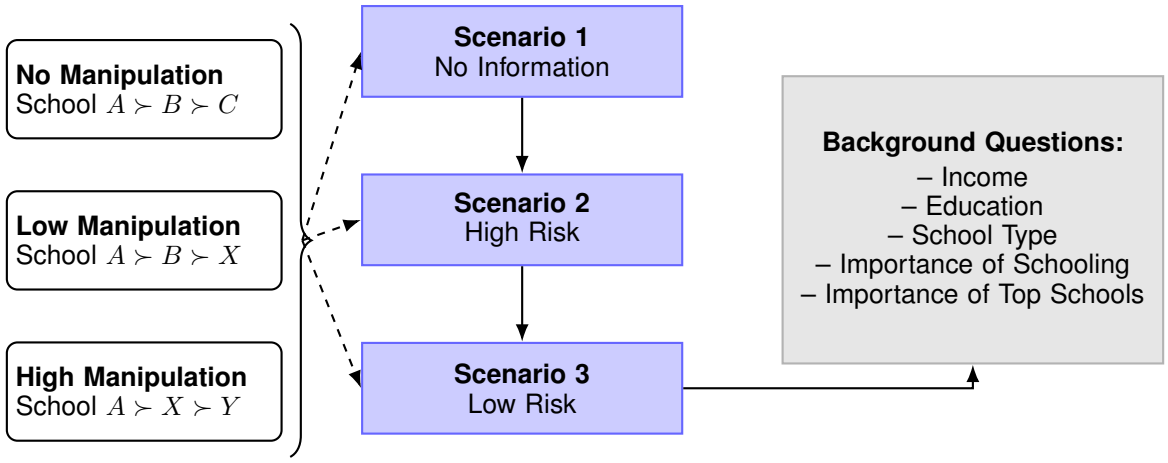


Figure 1: Survey structure and experimental design. Respondents select from one of three preference lists in each of three scenarios varying in the availability of information and strategic risk. Background questions follow the final scenario.

aspirational levels. In particular, we focus our investigation on the behavior of the higher SES parents, as the literature suggests that these are most prone to engage in strategic behavior.

Strategic Behavior in School Choice

Centralized school choice systems use algorithmic mechanisms to allocate students to schools based on stated preferences and school capacities. The Deferred Acceptance (DA) algorithm (Gale and Shapley 1962; Abdulkadiroğlu and Sönmez 2003) is widely used for its strategyproofness and stability guarantees. However, DA does not necessarily maximize welfare, leading to dissatisfaction among students assigned to lower-ranked choices (Abdulkadiroğlu et al. 2006). Alternative mechanisms such as Rank Minimization (RM) (Featherstone 2020; Ortega and Klein 2023) have been proposed to improve efficiency by minimizing the number of students receiving low-ranked school assignments, albeit at the cost of introducing incentives for strategic manipulation (Troyan 2022).

Empirical studies have documented that when mechanisms are manipulable, strategic behavior becomes widespread. Abdulkadiroğlu et al. (Abdulkadiroğlu et al. 2006) show manipulation under the original Boston mechanism, while Calsamiglia et al. (Calsamiglia, Haeringer, and Klijn 2010) report similar findings in Barcelona. Pathak and Sönmez (Pathak and Sönmez 2008) highlight that non-strategyproofness can disproportionately disadvantage less-informed families. Although these studies establish that manipulation occurs, they offer limited insight into *who* chooses to strategize and *how* individual differences affect strategic behavior.

Socioeconomic factors have been shown to influence school choice behaviors and strategic sophistication. Hastings and Weinstein (Hastings and Weinstein 2008) demonstrate that access to better information alters school selections across income groups, and He (He 2012) finds that higher-income families exhibit greater strategic behavior in school choice settings.

Beyond observational data, experimental and survey-based methods have been used to investigate strategic behavior in assignment systems. Bronfman et al. (Bronfman et al. 2018) conduct laboratory experiments to study strategic preference reporting under different mechanism designs, while Rees-Jones (Rees-Jones 2018) demonstrates that even with clear strategic incentives, participants often fail to optimally manipulate their preferences.

Behavioral Theories on Decision Making

Decision-making under uncertainty has long been studied in behavioral economics and psychology. The Security-Potential/ Aspiration (SP/A) theory, developed by Lopes and Oden (Lopes and Oden 1999), models individuals as balancing two competing motivations when facing risky choices: 1) Security, prioritizing the minimization of potential losses, and 2) Potential, focusing on maximizing the best possible outcomes. Aspiration levels serve as reference points in this trade-off, with individuals holding higher aspirations more willing to take calculated risks to achieve desired goals. SP/A theory explains probability distortions not through perceptual biases, as in Prospect Theory (Kahneman, Tversky et al. 1979), but through the emotional dynamics of hope and fear. It emerged in response to empirical violations of the independence axiom in Expected Utility Theory (Von Neumann and Morgenstern 1947), such as those demonstrated in the Allais and Ellsberg paradoxes (Keynes 2013). SP/A theory has been widely applied in financial decision-making and consumer behavior but has not been extensively explored in the context of school choice. Although expected utility theory models choice as maximizing expected value, it cannot fully explain the reluctance towards risking highly undesirable outcomes. Meanwhile, prospect theory adds probability weighting and loss aversion, but evaluates outcomes along a single utility scale. SP/A theory, by contrast, conceptually separates the drive to secure a satisfactory outcome from the aspiration to achieve

the best possible one, capturing the tendency for people to reject high-probability gains when they risk highly undesirable outcomes.

Our work contributes to this literature by designing a behavioral survey based on realistic Amsterdam school choice conditions, varying transparency of admission probabilities to observe how parents respond to strategic opportunities, and model this behavior using SP/A theory.

Methods

We conducted a quantitative behavioral survey (n = 140) to investigate how parental aspirations and socioeconomic background influence strategic behavior under a manipulable school choice mechanism.

We ground our survey design in the SP/A theory (Lopes and Oden 1999), as introduced in Section . This theory models decision-making under risk as a balance between potential, security, and aspiration. In the context of school choice, potential refers to the possibility of achieving a highly desirable outcome, such as being matched with a top-ranked school, while security reflects the desire to avoid the worst-case outcome, namely not being placed at any school at all. Aspiration represents the level of outcome parents hope to achieve, and the gap between this aspiration and their expected outcome influences their willingness to take risks.

Survey Design

Participants completed a hypothetical school choice task based on a manipulable matching mechanism. Each respondent was assigned a fixed true preference ranking over three fictional schools, $A \succ B \succ C$, and informed that it might be advantageous to misrepresent their preferences—e.g., by including oversubscribed schools such as X and Y that they did not genuinely prefer.

In each of three scenarios, participants were asked to select one preference ranking from three pre-constructed lists. The first list reflected truthful reporting ($A \succ B \succ C$). The second list, representing a *mild strategy*, included $A \succ B \succ X$, using the oversubscribed school X to improve the chance of placement in the top two choices. The third list, representing an *aggressive strategy*, prioritized two highly competitive schools ($A \succ X \succ Y$) to increase the likelihood of receiving the top choice through a riskier manipulation. The design of the strategies follows from the best-response analysis in (Tasnim et al. 2024), where it is shown that listing popular schools after one's top preferred schools can increase the likelihood of being assigned to those top schools. The mild strategy implements this by being truthful about the top two choices and placing one popular school later in the list, applying moderate pressure to secure a placement within the top two. The aggressive strategy is truthful only about the top choice and places two popular schools afterwards, exerting greater pressure to obtain the top-ranked school at the cost of a higher risk of non-placement.

The three scenarios varied in the degree of information available. In the *baseline* scenario, no admission probabilities were shown. In the second scenario, participants were shown realistic probabilities for each list, where strategic options carried substantial risk and offered limited benefit. In the third scenario, the probability of admission to the first-choice school on the aggressive list was artificially boosted to 90%, simulating a low-risk, high-reward condition. Probabilities in Scenarios 2 and 3 were generated using historical Amsterdam placement data under a manipulable rank-minimizing method similar to (Ortega and Klein 2023; Tasnim et al. 2024). Following each scenario, participants could optionally provide open-ended explanations for their choices. Additional demographic data were collected, including household income, education level, school track preference (e.g., pre-university vs. vocational), and perceptions of school quality and the importance of elite school access.

Variable	Category	Count (%)
Education	PhD	11 (7.9%)
	Master's degree	62 (44.3%)
	Bachelor's degree	32 (22.9%)
	Associate's degree	16 (11.4%)
	High School	18 (12.9%)
	Primary school	1 (0.7%)
Income	$\geq$ €100,000	49 (35.0%)
	€50,000–€100,000	44 (31.4%)
	€25,000–€50,000	18 (12.9%)
Track	Pre-University	43 (30.7%)
	General/Pre-University	29 (20.7%)
	General	21 (15.0%)
	Vocational	12 (8.6%)
	Practical education	31 (22.1%)

Table 1: Demographic characteristics of the sample ($n = 140$). Counts are rounded to the nearest whole number.

Hypothesis

We test the following hypotheses in our experiment:

H1—Strategic Behavior: When assignment probabilities are disclosed and the perceived risk of manipulation is low, parents are more likely to behave strategically.

H2—Strategy Type: Even under low-risk, most parents will avoid aggressive strategies and instead favor mild deviations.

H3—Risk Profile and Competition: Parents applying to more competitive school tracks will exhibit greater risk aversion, regardless of socioeconomic background.

H1 reflects how the disclosure of admission probabilities affects parents' perception of security, while H2 suggests that the perceived loss of security associated with aggressive strategies continues to dominate most parents' choices. H3 builds on the assumption that competitive school tracks elevate the perceived cost of failure. As a result, parents may prioritize security over potential when competition makes negative outcomes (such as being unplaced) more probable and consequential.

Sampling

Participants were reached through the Onderwijs Consumenten Organisatie (OCO), a non-profit organization supporting parents navigating the Amsterdam school system. Eligibility criteria included being a parent or guardian with a child eligible for Amsterdam's secondary school lottery. Participants were offered the opportunity to enter a prize lottery upon survey completion. The study was conducted under standards consistent with exemption from formal IRB review, given its minimal risk profile. The questionnaire for the survey is disclosed in the Appendix [1] .

Demographic information collected included parental income, education level, school aspiration attitudes, and prior experience with the school lottery process. The demographic of our sample is reported in Table 1.

Analysis

To evaluate the outlined hypotheses, we conducted a series of descriptive statistical tests. For H1, we compared participants' strategy choices across the three scenarios to assess whether transparency and reduced risk increase the incidence of strategic behavior. We applied McNemar's test (McNemar 1947) to paired responses from Scenario 2 and 3 to assess whether the observed decline in truthful reporting was statistically significant. For H2, we analyzed the distribution of non-truthful choices, comparing the prevalence of mild versus aggressive strategies across scenarios. For H3, we disaggregated strategy use by parent characteristics, particularly school track preference and aspiration level, and assessed whether behavioral differences across groups were consistent with increased risk aversion in competitive contexts. All comparisons were visualized and supported by chi-square tests of independence to identify statistically meaningful group differences.

Modeling Strategic Risk with SP/A Theory We estimated a discrete choice model grounded in the Security-Potential/Aspiration (SP/A) theory of decision-making under uncertainty. SP/A theory posits that individuals evaluate risky options based on two competing motivations: the *potential* to achieve an aspirational goal and the *security* of avoiding disappointing outcomes. Rather than assuming distortions in probability perception, as in Prospect Theory, SP/A models decision-making as an affective trade-off between hope for gain and fear of loss.

Formally, we define the utility of strategy s for parent i as:

$$U_{i,s} = \alpha_i \cdot \text{Potential}_s + (1 - \alpha_i) \cdot \text{Security}_s,$$

where $\alpha_i \in [0, 1]$ is a latent parameter capturing individual risk preference. Higher values of α_i indicate a stronger orientation toward potential (risk-seeking), whereas lower values reflect greater emphasis on security (risk aversion).

In our implementation, *potential* was defined as the probability of being placed at the top-ranked school in a given list, while *security* was defined as the probability of avoiding non-placement (i.e., being matched to any school). These probabilities were drawn directly from the scenario-specific admission estimates shown to participants during the survey.

To capture heterogeneity across individuals, we modeled α_i as a logistic function of standardized parent-level features:

$$\alpha_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \cdot \text{Education}_i + \beta_2 \cdot \text{Income}_i + \ldots)}}$$

We estimated a multinomial logit model over the three strategic choices (truthful, mild, aggressive) in each scenario, with utilities computed according to the SP/A formulation above. A single model was fit jointly across all three scenarios, using the scenario-specific potential and security values for each strategy. Each participant contributed one observation per scenario; standard errors were clustered at the respondent level to account for within-subject correlation. Although SP/A theory can be extended with non-linear specifications such as aspiration cutoffs or weighted utility distortions, we employed a linear formulation to prioritize model interpretability and ensure stable parameter estimation given our small sample size of 140.

For comparison, we implemented two benchmark models: an Expected Utility (EU) model (Von Neumann and Morgenstern 1947) with linear probability-weighted outcomes, and a Prospect Theory (PT) model (Kahneman, Tversky et al. 1979) using a single-parameter Prelec weighting function. All models were estimated via maximum likelihood using default convergence settings. Model performance was evaluated using log-likelihood, Akaike Information Criterion (AIC), and predictive accuracy on held-out data.

Results

Descriptive Patterns of Strategic Behavior

Participants' strategy choices varied systematically across the three experimental conditions, as shown in Figure 2. In the baseline scenario, where no information about admission probabilities was provided, 80.1% of participants submitted the truthful list. This share declined to 57.4% under the high-risk transparency condition and further to 40.4% when the aggressive strategy was presented as favorable. McNemar's test confirms that the reduction in truth-telling between the High Risk and Low Risk conditions is statistically significant ($\chi^2 = 14.69$, $p = 0.00013$). These results support our first hypothesis (H1), indicating that transparency combined with low perceived risk increases the incidence of strategic behavior.

Scenarios	Truthful (%)	Mild (%)	Aggressive (%)
Baseline	80.14	14.18	5.67
High Risk	57.45	38.30	4.26
Low Risk	40.43	20.57	39.01

Table 2: Strategy choices across the three experimental scenarios.

[1] Additional technical appendices and supplementary analyses are available at: https://doi.org/10.5281/zenodo.16811373.

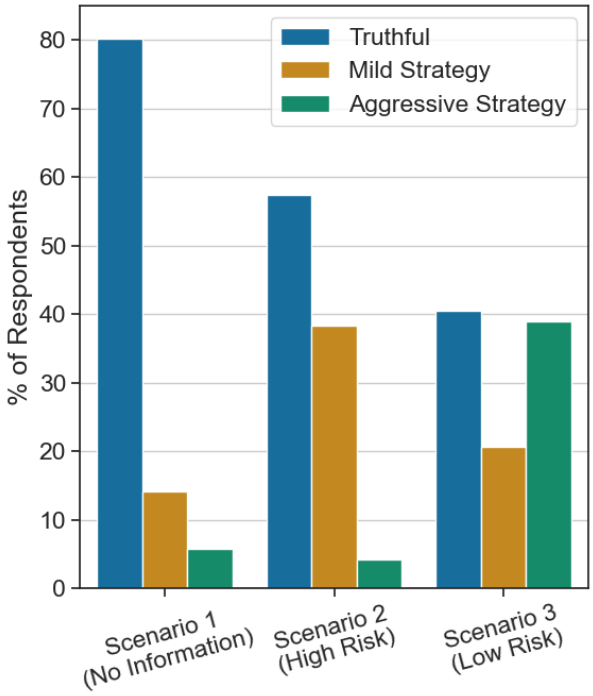


Figure 2: Strategic behavior by scenario. Each bar shows the percentage of participants choosing Truthful, Mild Strategy, or Aggressive Strategy across Baseline, High Risk, and Low Risk scenarios.

We also examine the specific types of strategic behavior that participants adopted. As shown in Table 2, the share of participants choosing the aggressive strategy (List 3) increased sharply from 4.3% in the High Risk condition to 39.0% in the Low Risk condition. This shift suggests that participants were responsive to the underlying probability structure and that some were willing to pursue high-reward strategies when the risk of non-assignment was reduced.

However, even when the aggressive strategy offered a 90% chance of securing the top choice, many participants ($\sim$60%) continued to avoid it. Among strategic respondents, a sizable proportion favored the mild deviation (List 2), which carried less risk of being misplaced. In Scenario 3, 20.6% of all participants selected the mildly strategic option, despite the more favorable probabilities associated with aggressive behavior. This sustained preference for lower-risk strategies provides support for our second hypothesis (H2): while transparency shifts behavior away from truth-telling, risk aversion continues to moderate the extent of strategic deviation.

Risk Preference Estimation and Model Fit

The SP/A-based discrete choice model provided a substantially better fit to the data than a constant-risk baseline. A likelihood ratio test comparing the full model to the baseline yielded $\chi^2 = 258.55$ with $p < 10^{-50}$. The model also achieved a pseudo-R^2 of 0.276, and outperformed the

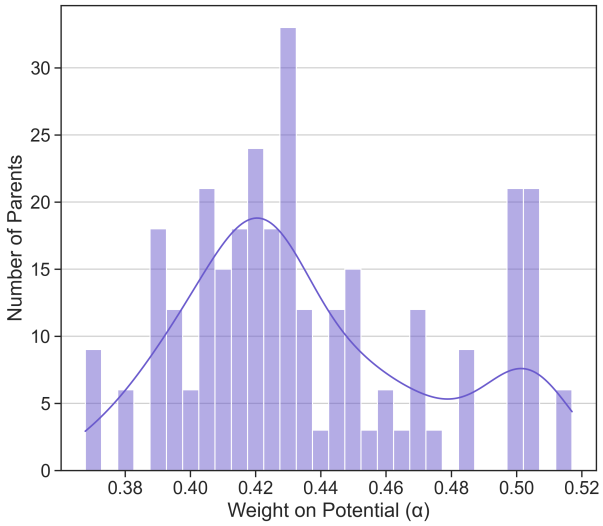


Figure 3: Distribution of fitted risk preference parameters α_i across respondents. Higher α indicates greater emphasis on potential vs. security.

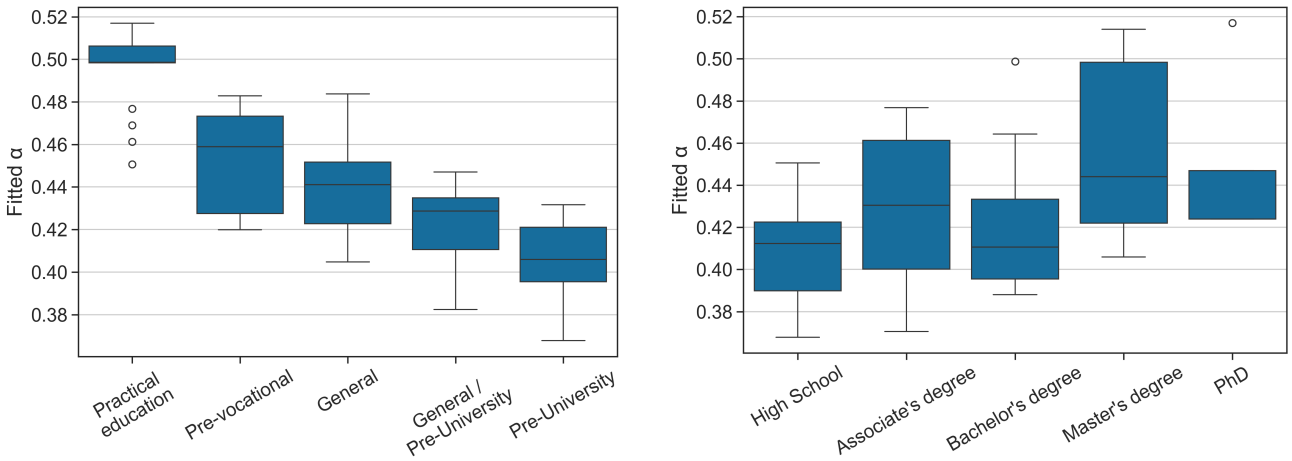
baseline on model selection criteria (AIC: 689.1 vs. 937.6; BIC: 713.4 vs. 941.7). While none of the individual covariates reached conventional levels of statistical significance ($p > 0.05$), the overall model captures meaningful variation in strategy behavior, suggesting that the combination of features helps explain latent risk preferences.

Figure 3 shows the estimated distribution of individual risk parameters α_i. The distribution is bimodal, with one group of participants placing substantial weight on security (e.g., $\alpha_i < 0.45$), and another group moderately prioritizing potential outcomes (e.g., $\alpha_i > 0.45$). Very few respondents exhibit strongly risk-seeking behavior ($\alpha_i > 0.6$), even under favorable conditions, consistent with the observed rarity of aggressive strategic choices.

We next examined which parent characteristics best explain variation in α_i. As shown in Table 3 and Figure 4a, school track preference emerged as a key predictor: parents whose children were applying to competitive programs tended to exhibit greater risk aversion, possibly reflecting heightened sensitivity to the consequences of poor placement outcomes. Education was weakly positively associated with risk-seeking, while income and attitudinal factors, such as the importance placed on elite schooling, showed no significant effects once other variables were accounted for.

These findings provide partial support for our third hypothesis (H3). While parents do become more strategic under low-risk conditions, their choices appear moderated by how they perceive and manage risk, not by education or income.

There are several plausible explanations for the lack of statistically significant individual predictors despite strong overall model fit. Firstly, the sample size ($n = 140$) may limit statistical power to detect modest effects. While correlations among covariates such as income, education, and



(a) School Type

(b) Parent Education

Figure 4: Distribution of fitted risk preference parameters α_i across demographic variables of respondents.

Covariate	**Estimate** (β)	**Effect on** α_i
Intercept	0.058	Baseline
Education (std.)	0.088	↑ Risk-seeking
Income (std.)	–0.023	No effect
School Type	–0.094	↓ Risk-seeking
Importance: Schooling	–0.020	Negligible
Importance: Top School	–0.00006	Negligible

Table 3: Estimated coefficients for covariates explaining heterogeneity in α_i.

Model	**Log-Likelihood**	**AIC**	**BIC**	**Pseudo-R^2**
SP/A	338.55	665.09	647.40	1.73
EU	-464.71	931.43	934.37	0.00
PT	-464.71	933.43	939.32	≈ 0.00

Table 4: Model comparison using log-likelihood, AIC, BIC, and pseudo-R^2 for SP/A, Expected Utility (EU) and Prospect Theory (PT) models.

aspiration are moderate, multicollinearity does not appear to be a primary concern (see Appendix). Alternatively, the influence of these features may be context-dependent or mediated by unobserved factors, requiring further study to disentangle their role in shaping risk preferences in school choice settings.

We also compared the SP/A model against two baselines: Expected Utility (EU) and Prospect Theory (PT). Table 4 reports model fit statistics across specifications. The SP/A model achieves the highest log-likelihood (338.55), outperforming both EU and PT, which converge to -464.71. AIC, BIC, and pseudo-R^2 metrics likewise favor the SP/A model, indicating better in-sample fit.

Discussion

Our research centers on examining the behavior of parents in school choice mechanisms that are not strategyproof, and under varying levels of risk. We observe that a considerable share of respondents chose to report their preferences truthfully, even when given information that made the risks and potential gains of strategic behavior transparent.

However, the statistically significant increase in strategic behavior from Scenario 2 to Scenario 3 suggests that parents are sensitive to changes in the strategic landscape, particularly when the cost of deviation (i.e. risk) is lowered. Indeed, the preference for List 2 over the more aggressive List 3, even under favorable conditions, indicates that risk aversion plays a central role in shaping behavior and exceeds the importance of reward. These patterns align with the qualitative predictions of SP/A theory: participants tend to adopt secure strategies when uncertainty is high and only transition toward potential-maximizing strategies when the environment becomes less risky.

While strategic behavior does increase when probabilities are disclosed, its uptake appears to be governed not merely by access to information but also by individuals' tolerance for risk which, crucially, is not uniform. High-SES parents, often presumed to be the primary beneficiaries of manipulable systems, did not disproportionately adopt aggressive strategies in our study. Instead, parents at the most competitive school levels showed higher levels of risk aversion. We do not find strong evidence that income or education systematically explains variation in strategic behavior.

These findings should be interpreted in light of key design constraints. In order to reach parents currently participating in the school choice process, we recruited participants through informational seminars for the ongoing placement year. To avoid compounding their cognitive burden, the survey was designed to be short and completable within five minutes. Many participants were already navigating a com-

plex matching system and facing emotionally charged decisions; the idea of deliberately crafting a strategic submission added a further layer of stress that may have shaped both engagement and responses.

This difficulty in forming and executing strategies under uncertainty underscores the need for a behavioral model that goes beyond rational choice. The SP/A model provides a compelling explanation for this. Rather than treating parents as either fully rational or systematically biased, SP/A theory captures a more affective mode of decision-making, where aspirations are weighed against perceived threats. The model's better fit relative to Expected Utility and Prospect Theory supports the claim that this balancing act governs how families engage with school choice mechanisms. Importantly, the model reveals a bimodal distribution of risk preferences, suggesting that population-level behavior cannot be adequately understood through average tendencies alone.

From a policy perspective, these findings challenge the notion that limiting transparency is a necessary safeguard against strategic behavior. Rather than concealing information, school choice platforms might better support fairness by disclosing probabilities alongside tools that communicate the risks of manipulation in intuitive and context-sensitive ways. Designing for transparency need not mean encouraging exploitation: our findings suggest parents are highly risk-averse, and only shift to aggressive strategies when the perceived chance of success is exceedingly high—in our case, 90%. Prior simulation studies in Amsterdam suggest that strategic reporting improves outcomes for only 1–2% of students (Tasnim et al. 2024), making this level of confidence in strategy success extremely unrealistic. This gap between perceived and actual risk highlights the opportunity to build interactive tools that display the probabilities of admission under different ranking strategies, explicitly including uncertainty and known error bounds. Such interventions may deter manipulation not by limiting information, but by making risk legible.

These findings should be interpreted in light of several limitations. Our results are based on responses to hypothetical scenarios, and real-world implementation involving more students and schools could involve additional dynamics. Nonetheless, the consistent tendency of parents to minimize the risk of non-placement suggests that these patterns may generalize in live settings where the consequences directly affect their children. The simplified three-school setup allowed us to isolate strategic risk preferences, but actual school choice involves more complex preferences, repeated decisions, heterogeneous stakes, and social learning. Future work should test similar probability disclosures in full-scale choice environments, ideally in collaboration with school districts, to assess how these patterns translate when participants have diverse preferences. The small sample size of low-SES respondents also constrains our ability to draw strong subgroup conclusions. Furthermore, because our sample is drawn from Amsterdam, results may reflect sociocultural norms specific to the Netherlands. School choice policies, however, are often adapted internationally, suggesting similar dynamics could emerge elsewhere. Future work should extend this study to multiple cities and policy settings to examine how institutional context, cultural factors, and mechanism design interact in shaping strategic behavior.

Overall, this work underscores the value of integrating behavioral theory into the evaluation of algorithmic systems, particularly in high-stakes domains like education where institutional trust and individual aspiration intersect. From a policy perspective, our findings suggest that the trade-off between high efficiency and manipulability of an algorithm may not be as sharp as previously thought, and that transparency, rather than fueling strategic behavior, may in fact mitigate it. This contributes to the broader debate on transparency in human–AI interactions, showing that its effects depend on context, the nature of the information disclosed, and user incentives. Furthermore, given the nexus between educational opportunity and equality, what merits further investigation is not only the use of AI in education, but also its application in the broader educational ecosystem, including school choice.

Conclusion

In this study, we investigated how probability transparency in school choice mechanisms influences parental strategic behavior. Using a survey of 140 parents in Amsterdam, we examined whether aspiration levels and socioeconomic background affect the tendency to manipulate school preferences under a non-strategyproof allocation method. While transparency increases strategic behavior, a large fraction of parents still choose truth-telling. Among those who deviate, most opt for mild, low-risk strategies rather than aggressive manipulation. To interpret these patterns, we modeled parental decision-making using the Security-Potential/Aspiration (SP/A) framework, which captures the affective trade-off between avoiding bad outcomes and striving for aspirational goals. The SP/A model not only outperforms Expected Utility and Prospect Theory baselines, but also reveals a bimodal distribution of risk preferences—underscoring that population-level behavior cannot be explained by average tendencies alone. These findings have direct implications for the design of algorithmic school choice mechanisms, particularly when considering non-strategyproof methods that allow gaming. Rather than restricting transparency, policymakers should focus on communicating the risks of strategic behavior clearly and intuitively, enabling parents to make informed decisions. Additionally, school choice models should account for heterogeneous behavioral responses, as a uniform transparency policy may not work equally well for all families. Future work could explore behavioral modeling techniques such as agent-based simulations or reinforcement learning to understand better how parents interact with different allocation methods over time. Further validation through controlled experiments and comparative studies of non-strategyproof mechanisms would enhance our understanding of strategic behavior in realistic settings.

Ethics Statement

This study was conducted in close collaboration with local education policymakers and community partners involved in school assignment in the city of Amsterdam. All survey participants were recruited on a voluntary basis and provided informed consent prior to participation. The survey was anonymous, and no personally identifiable information was collected. The study was considered exempt from review by the institutional review board. The research was designed to minimize any potential risks to participants, and the framing of hypothetical school choice scenarios was carefully piloted to avoid distress or misinformation. Participants were debriefed about the nature and purpose of the study. We acknowledge that any algorithmic or mechanism-design intervention in school allocation systems must be evaluated in light of local social, economic, and political contexts. We do not propose any specific intervention as universally optimal. Rather, we aim to provide empirical insight into behavioral patterns that may inform context-sensitive and equity-oriented policy decisions.

Acknowledgments

This research was supported by the Innovation Center for AI (ICAI, The Netherlands) and the City of Amsterdam. We are deeply grateful to Floor Kaspers and Lidewij Koren from the Onderwijs Consumenten Organisatie (OCO) for their tireless support and constructive feedback in shaping and facilitating this study. We also thank Nicolas Choquette-Levy, Elisabeth H. Krueger, Fernando P. Santos, and Arjan Vreeken for their insightful advice and feedback throughout the development of this work. Finally, we extend our sincere appreciation to the schools and parents who generously participated in this study, making this research possible.

References

Abdulkadiroğlu, A.; Pathak, P.; and Roth, A. 2009. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match. *American Economic Review*, 99(5): 1954–78.

Abdulkadiroğlu, A.; Pathak, P.; Roth, A.; and Sönmez, T. 2006. Changing the Boston school choice mechanism.

Abdulkadiroğlu, A.; and Sönmez, T. 2003. School choice: A mechanism design approach. *American economic review*, 93(3): 729–747.

Abdulkadiroğlu, A.; and Sönmez, T. 1998. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3): 689–701.

Aksoy, S.; Azzam, A.; Coppersmith, C.; Glass, J.; Karaali, G.; Zhao, X.; and Zhu, X. 2013. School choice as a one-sided matching problem: Cardinal utilities and optimization. *arXiv preprint arXiv:1304.7413*.

Bogomolnaia, A.; and Moulin, H. 2001. A new solution to the random assignment problem. *Journal of Economic theory*, 100(2): 295–328.

Bronfman, S.; Alon, N.; Hassidim, A.; and Romm, A. 2018. Redesigning the israeli medical internship match. *ACM Transactions on Economics and Computation (TEAC)*, 6(3-4): 1–18.

Calsamiglia, C.; Haeringer, G.; and Klijn, F. 2010. Constrained school choice: An experimental study. *American Economic Review*, 100(4): 1860–74.

Featherstone, C. R. 2020. Rank efficiency: Modeling a common policymaker objective. *Unpublished paper, The Wharton School, University of Pennsylvania.[25, 28, 45]*.

Gale, D.; and Shapley, L. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1): 9–15.

Hastings, J. S.; Van Weelden, R.; and Weinstein, J. M. 2007. Preferences, information, and parental choice behavior in public school choice.

Hastings, J. S.; and Weinstein, J. M. 2008. Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly journal of economics*, 123(4): 1373–1414.

He, Y. 2012. Gaming the Boston School Choice Mechanism in Beijing. Toulouse School of Economics Working Paper.

Hulshof, M. 2025. Opinie: 'Ineens mogen kinderen niet weten hoe groot hun toelatingskans op een scholen is – dat is onmisbare informatie'.

Kahneman, D.; Tversky, A.; et al. 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2): 363–391.

Keynes, J. M. 2013. *A treatise on probability*. Courier Corporation.

Khaddari, R. 2022. Weer minder achtstegroepers kunnen naar eerste voorkeursschool: 'Volgend jaar wordt het beter'.

Lopes, L. L.; and Oden, G. C. 1999. The role of aspiration level in risky choice: A comparison of cumulative prospect theory and SP/A theory. *Journal of mathematical psychology*, 43(2): 286–313.

McNemar, Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2): 153–157.

Ortega, J.; and Klein, T. 2023. The cost of strategy-proofness in school choice. *Games and Economic Behavior*.

Pais, J.; and Pintér, Á. 2008. School choice and information: An experimental study on matching mechanisms. *Games and Economic Behavior*, 64(1): 303–328.

Parliament, E.; and Council, E. 2022. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance). *OJ*.

Parliament, E.; and Council, E. 2023. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). *OJ*.

Pathak, P.; and Sönmez, T. 2008. Leveling the playing field: Sincere and sophisticated players in the Boston mechanism. *American Economic Review*, 98(4): 1636–52.

Peeters, R. 2020. The agency of algorithms: Understanding human-algorithm interaction in administrative decision-making. *Information Polity*, 25(4): 507–522.

Rees-Jones, A. 2018. Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match. *Games and Economic Behavior*, 108: 317–330.

Tasnim, M.; Weesie, Y.; Ghebreab, S.; and Baak, M. 2024. Strategic manipulation of preferences in the rank minimization mechanism. *Autonomous Agents and Multi-Agent Systems*, 38(2): 44.

Troyan, P. 2022. Non-obvious manipulability of the rank-minimizing mechanism. *arXiv preprint arXiv:2206.11359*.

Ustun, B.; Spangher, A.; and Liu, Y. 2019. Actionable Recourse in Linear Classification. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, 10–19. New York, NY, USA: Association for Computing Machinery. ISBN 9781450361255.

Von Neumann, J.; and Morgenstern, O. 1947. *Theory of games and economic behavior, 2nd rev*. Princeton university press.

Wang, Q.; Huang, Y.; Jasin, S.; and Singh, P. V. 2023. Algorithmic transparency with strategic users. *Management Science*, 69(4): 2297–2317.

Zerilli, J.; Bhatt, U.; and Weller, A. 2022. How transparency modulates trust in artificial intelligence. *Patterns*, 3(4).