# Modeling and Solving Stable Matching under Probabilistic Preferences with Large Language Models

**Yuqi Kong[1*], Shiyu Liu[2*], Jiaxu Li[3], Hongtao Liu[1], Qi Qi[1], Weiran Shen[1†],**

[1]Gaoling School of Artificial Intelligence, Renmin University of China,
[2]School of Information, Renmin University of China,
[3]Faculty of Science, National University of Singapore
**Correspondence:** {kongyuqi, liushiyu9, shenweiran}@ruc.edu.cn

## Abstract

Large language models (LLMs) have recently demonstrated strong capabilities in understanding and simulating human decisions, suggesting a new way to use LLMs as tools to study social systems. We study two-sided-matching markets, such as dating and job matching. Classical matching models assume deterministic, strict preferences, which violate the real-world setting. We focus on stable matching under stochastic decision behavior and use LLMs to simulate human-like preferences and probabilistic choice patterns. Based on this, we introduce Expected Blocking Pairs (EBP), a continuous measure to quantify stability that generalizes the classic blocking pair notion. We further propose a Hybrid GS–LLM matching method that integrates the celebrated Gale–Shapley (GS) algorithm with probabilistic acceptance decisions. Experiments show that the proposed hybrid method outperforms classical baselines in terms of stability, suggesting that LLMs provide a principled tool for modeling human decisions and for improving robustness of matching under uncertainty.

## 1 Introduction

The rapid development of LLMs has led to increasingly strong evidence that these models exhibit human-like decision behaviors. Trained on large-scale human-generated data, LLMs naturally inherit cognitive characteristics such as ambiguity, bias, and bounded rationality. These properties make LLMs not only powerful predictive models, but also promising tools for studying and evaluating complex social systems. Recent studies suggest that LLMs serve as a promising tool for modeling stochastic human decision-making and analyzing the decision-driven social mechanisms in multi-agent settings (Acerbi and Stubbersfield, 2023; Strachan et al., 2024; Cheung et al., 2025). For instance, LLMs are already used to reproduce classic phenomena from behavioral economics, such as endowment effects and loss aversion, while simultaneously displaying systematic deviations from full rationality (Horton, 2023; Liu et al., 2025; Xie et al., 2025; Capraro et al., 2025).

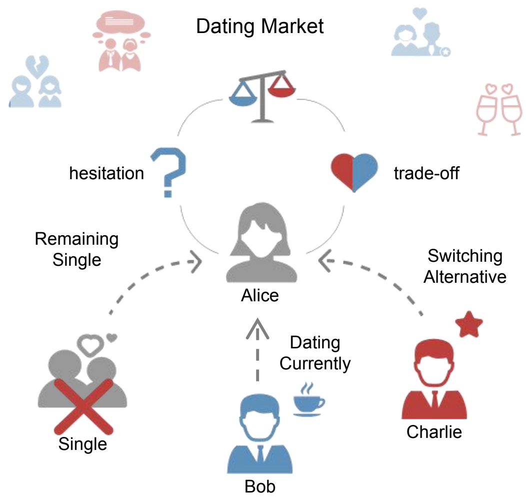



Figure 1: **Decision-making under probabilistic preferences in a dating market.** An agent (Alice) chooses whether to stay with her current match (Bob), switch to an alternative (Charlie), or remain single under bounded rationality. Such stochastic choices motivate evaluating stability in expectation.

Many real-world scenarios involve two-sided matching. Prominent examples include dating and marriage markets, where individuals decide whether to start, maintain, or terminate relationships, as well as job markets and school admissions. To study such environments, a foundational framework in economics and computer science is the two-sided matching problem. The stable marriage problem (SMP) (Gale and Shapley, 1962) is, perhaps, the most well-known and widely applied.

In the classical formulation of SMP, two disjoint sets of agents, $W$ and $M$, are given, where each agent in $W$ has a strict total preference order over all agents in $M$, and vice versa. The goal is to

*Equal contribution.
†Corresponding author.

find a matching, i.e., a one-to-one correspondence between $W$ and $M$, such that the outcome contains no blocking pair, namely, no pair of agents who would both prefer to be matched with each other rather than with their assigned partners under the current matching. A matching that admits no blocking pair is referred to as a stable matching. In their seminal paper, Gale and Shapley also proposed an algorithm, known as the Gale-Shapley algorithm, which is guaranteed to find a stable matching with no blocking pairs. The stable matching theory and its applications have played a central role in two-sided matching market design and were formally recognized by the 2012 Nobel Memorial Prize in Economic Sciences.

The classical SMP assumes that each agent has a strict, unchanging preference order and will always choose a higher-ranked partner over a lower-ranked one. However, these assumptions fail to capture the behaviors of real-world matching markets. In practice, preferences are noisy, multi-dimensional, and influenced by switching costs and contextual factors. As a result, agents do not always switch to a more preferred alternative deterministically. Instead, partner-switching decisions are inherently probabilistic. Under such conditions, the classical notion of stability becomes overly idealized: it neither captures stochastic decision behavior nor quantifies how stable a matching is under preference uncertainty.

Based on these observations, a natural question is: given LLMs' capability to simulate human-like decisions, can LLMs be used to assist with matching problems, for instance by using them to model probabilistic preferences in matching markets, and can LLMs help find a more stable matching under such settings?

Motivated by these considerations, we make the following contributions:

- We demonstrate that the decision patterns of LLMs exhibit strong consistency with the Quantal Response (QR) models of bounded rationality in behavioral economics (McKelvey and Palfrey, 1995). Using LLM agents as stochastic decision-makers, we estimate switching behavior that provides an interpretable probabilistic abstraction of preference-driven decisions.

- We introduce Expected Blocking Pairs (EBP), a continuous stability metric that aggregates the probability that each potential pair forms a blocking pair. This extends the classical binary notion of stability to a measurable and comparable quantity.

- We propose a hybrid method which we call the Hybrid GS-LLM Matching method and conduct experiments on a real speed-dating dataset (Fisman et al., 2006). Experiments show that this hybrid method improves the stability compared to the classical Gale-Shapley algorithm.

## 2 Related Work

Research related to our work falls broadly into two categories. The first line of research studies classical stable matching and its extensions under preference uncertainty and relaxed notions of stability. The second line investigates the use of large language models as agents for modeling human preferences and decision-making. Our work connects these two strands by introducing probabilistic stability notions induced by LLM-based agents.

### 2.1 Classical Stable Matching and Stability

The stable marriage problem, introduced by Gale and Shapley (Gale and Shapley, 1962), is a cornerstone of matching theory and has been extensively studied in economics and market design (Roth and Sotomayor, 1992). In this framework, agents on both sides are assumed to hold complete and deterministic preference rankings, and stability is defined as the absence of any blocking pair. These models have also been extended to many-to-one settings such as college admissions and residency matching (Ngoc et al., 2024).

Motivated by the gap between deterministic theoretical models and noisy real-world preferences, a growing body of work incorporates uncertainty into preference modeling. Examples include uncertain linear preferences (Aziz et al., 2020), multi-modal preference representations (Wen et al., 2022; Yang et al., 2025b), and multi-layer preference (Bentert et al., 2023).

This line of work also considers relaxed notions of stability, where a matching is allowed to admit a limited number of blocking pairs or is only approximately stable. Representative examples include the Minimum Blocking Pairs Matching (Gao et al., 2025), analyses of the combinatorial structure of the stable matching lattice (Alkan and Yildiz, 2023), as well as mechanisms that tolerate limited

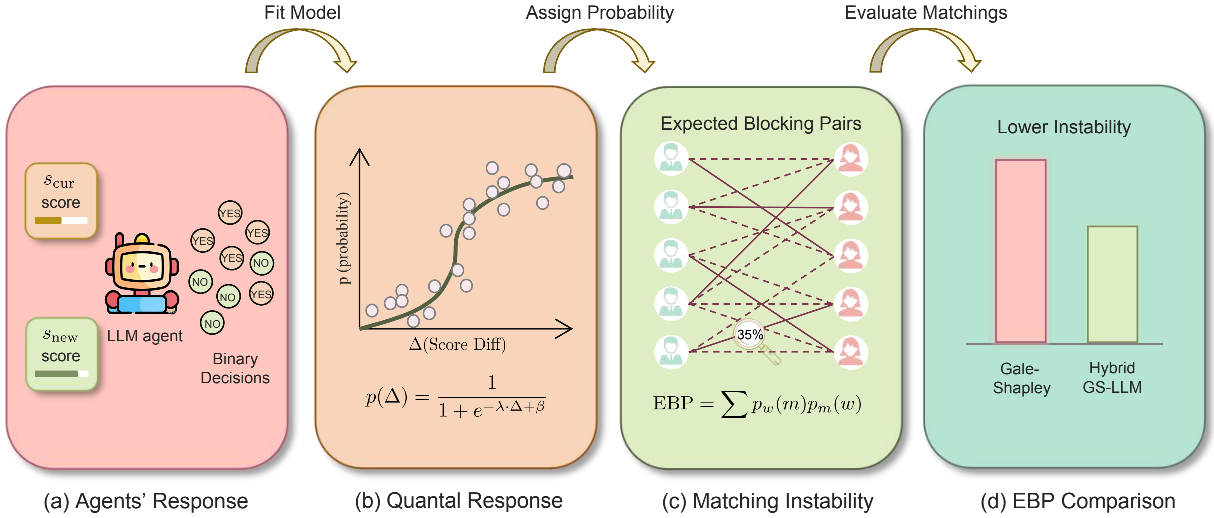



Figure 2: **Overview of the proposed LLM-based probabilistic matching framework**. (a) Given score-based evaluations of partners, LLM agents generate stochastic binary decisions. (b) The decisions are summarized using a quantal response model, which maps score differences to probabilistic switching behavior. (c) The resulting probabilistic preferences are used to compute Expected Blocking Pairs. (d) Classical Gale–Shapley and the proposed Hybrid GS–LLM methods are evaluated under the EBP metric

violations of stability (Ashlagi et al., 2025). Previous work also considers probabilistic preference models, where preferences are drawn from underlying distributions (Pittel, 2025); however, stability is still evaluated deterministically once preferences are realized.

In contrast to these approaches, existing studies do not explicitly characterize the blocking pairs as stochastic events with well-defined probabilities, nor do they provide a quantitative measure of stability that captures the expected extent of instability induced by noisy decision behavior.

### 2.2 LLM-based Modeling of Agents

Recent advances in LLMs have opened new possibilities for modeling human preference and decision-making. Studies have shown that human-like object concept representations naturally emerge in multi-modal LLMs (Du et al., 2025). Previous work has also discussed whether LLM-based agents can reproduce classical behavior economic phenomena (Horton, 2023; Lorè and Heydari, 2024; Liu et al., 2025), and whether they can be embedded into multi-agent simulations with memory, planning, and interaction (Park et al., 2023; Ferraro et al., 2024).

Although recent work has evaluated LLMs as algorithmic reasoners capable of executing matching mechanisms given ranked preferences (Liu et al., 2025), there remains a notable gap in treating LLMs as stochastic decision agents that generate noisy, probabilistic preferences. In particular, the implications of such probabilistic decision behavior for stability in matching markets remain largely unexplored.

## 3 Model

In this section, we introduce a probabilistic modeling framework for two-sided matching under preference uncertainty.

We generalize the classical two-sided matching framework by relaxing its deterministic behavioral assumptions while preserving its core structure. In the classical framework, each agent holds a deterministic preference over candidates, which is a complete preference list of agents on the other side. The celebrated Gale-Shapley algorithm iteratively produces a stable matching outcome through a simple proposal process: unmatched (“single”) agents propose to their most preferred remaining candidates, receivers become tentatively matched (“engaged”) to the best offer so far and reject the rest, and rejected agents become unmatched again and continue proposing until no possible proposals remain.

In our setting, these preferences are grounded in scalar compatibility scores, where a higher score implies a stronger preference. For instance, agent $i$ is given three candidates with scores $s_i(a) = 0.9$, $s_i(b) = 0.7$, and $s_i(c) = 0.2$, then the agent $i$'s preference order is $a \succ b \succ c$. We then relax the classical setting by modeling partner-choice behavior with stochastic and irrational agents. Based on this model, we define a quantitative stability metric and propose a hybrid matching algorithm.

### 3.1 Quantal Response Modeling of Partner Switching

We model probabilistic partner-switching behavior using the quantal response framework from behavioral economics, which is a standard framework in behavioral economics and game theory. This model posits that agents are more likely to choose higher-utility options, but may still choose inferior alternatives with a nonzero probability.

We represent the evaluation of agent $i$ toward partner $j$ under a given LLM model $a$ by a scalar score $s_i^a(j)$. Utilities are defined as an affine transformation

$$u_i^a(j) = \lambda^a s_i^a(j) + \alpha^a,$$

where $\lambda^a > 0$ controls the sensitivity of model $a$ to utility differences.

Under the quantal response model, when choosing an option from a finite set $\mathcal{M}$, the probability of selecting option $j$ is

$$\Pr(j) = \frac{e^{u_i^a(j)}}{\sum_{k \in \mathcal{M}} e^{u_i^a(k)}}.$$

In the binary choice setting relevant to partner switching, when agent $i$ compares two candidates $j$ and $k$, the probability of choosing $k$ is

$$\begin{aligned}\Pr(k) &= \frac{e^{u_i^a(k)}}{e^{u_i^a(j)} + e^{u_i^a(k)}} \\ &= \frac{1}{1 + e^{u_i^a(j) - u_i^a(k)}} \\ &= \frac{1}{1 + e^{-\lambda^a\left(s_i^a(k) - s_i^a(j)\right)}}.\end{aligned}$$

When agent $i$ is already matched with partner $j$ and considers switching to a new partner $k$, we introduce an offset parameter $\beta^a$ to capture the switching cost:

$$\Pr\left(E_{i:j\to k}^a\right) = \frac{1}{1 + e^{-\lambda^a\left(s_i^a(k) - s_i^a(j)\right) + \beta^a}}, \quad (1)$$

where $E_{i:j\to k}^a$ denotes the event that agent $i$ switches from $j$ to $k$ under LLM model $a$.

### 3.2 Stability Measure

Classical stability in two-sided matching is a binary notion: a matching is either stable or unstable. Under probabilistic decision behavior, stability should instead be evaluated in expectation, requiring a quantitative measure that captures the likelihood of mutually beneficial deviations.

Let $A$ be a matching, and $\mu_A(i)$ the partner assigned to agent $i$ under $A$, where $\mu_A(i) = i$ indicates that $i$ remains single. For any candidate $j$, define the score difference under LLM model $a$ as

$$\Delta_i^a(j) = s_i^a(j) - s_i^a(\mu_A(i)).$$

Under the quantal-response-based switching model, the probability that agent $i$ deviates from $\mu_A(i)$ to $j$ is

$$\begin{aligned}p_i^a(j) &= \Pr\left(E_{i:\mu_A(i)\to j}^a\right) \\ &= \frac{1}{1 + e^{-\lambda^a \Delta_i^a(j) + \beta^a}}.\end{aligned}$$

**Definition 3.1** (Expected Blocking Pairs)**.** Given a matching $A$, the Expected Blocking Pairs (EBP) is defined as

$$\mathrm{EBP}(A) = \sum_{(m,w)\notin A} p_m^a(w)\, p_w^a(m),$$

where the summation ranges over all unmatched cross-set pairs.

The term $p_m^a(w)p_w^a(m)$ represents the probability that agents $m$ and $w$ would both prefer to deviate and match with each other, assuming independent switching decisions.

*Remark* 3.2. Our EBP measure generalizes the classical notion of blocking pair counts, i.e., the number of blocking pairs. When switching becomes deterministic, i.e., $p_i^a(j) \in \{0, 1\}$ for all $i, j$, $\mathrm{EBP}(A)$ reduces exactly to the classical count of blocking pairs.

EBP provides a continuous measure of instability under stochastic decision behavior: lower values indicate that mutually beneficial deviations are unlikely in expectation.

### 3.3 Hybrid GS-LLM Matching Method

The classical Gale-Shapley algorithm operates purely on ordinal preference rankings and guarantees stability only in a deterministic sense. It is not designed for probabilistic metrics like EBP. Applying the Gale-Shapley algorithm directly under a probabilistic setting may lead to suboptimal

**Algorithm 1** Hybrid GS-LLM Matching

**Input:** Proposers $M$, receivers $W$; score-based preference orders; probabilistic switching model $p(\cdot)$

**Output:** Matching $A$ (with `single` as a valid outside option)

```
A ← ∅
while there exists a single proposer m ∈ M who has remaining candidates do
    m proposes to remaining highest-ranked w
    if w is single then
        w accepts m according to p_w(m)
        if accepted then A ← A ∪ {(m, w)}
    else
        Let (m′, w) ∈ A be w's current match
        w switches to m according to p_w(m)
        if switched then
            A ← (A \ {(m′, w)}) ∪ {(m, w)}
return A
```

outcomes when stability is evaluated in expectation. This motivates a hybrid mechanism that integrates the structure of Gale-Shapley with LLM-based decision models.

Our Hybrid GS-LLM Matching method preserves the proposal-based structure of the Gale-Shapley algorithm, while replacing the deterministic preference comparison with an LLM-based probabilistic acceptance decision. Upon receiving a proposal, an agent decides whether to accept it probabilistically. The matching gets updated with accepted proposals, and rejected agents continue proposing. The process terminates when no further proposals are possible. Our analysis prioritizes the stability properties induced by these agents, distinguishing our work from evaluations of behavioral fidelity.

## 4 Experiments

In this section, we empirically evaluate the proposed probabilistic matching framework. We first validate whether LLMs produce preference judgments that are meaningfully aligned with human behavior using a speed-dating dataset. We then examine whether LLM-induced switching behavior follows a quantal response pattern. Finally, we compare different matching mechanisms under uncertainty using the expected blocking pairs as a stability criterion.

### 4.1 Experimental Setup

**Dataset.** We use the speed-dating dataset collected by Columbia Business School, which was gathered in a series of speed-dating experiments conducted between 2002 and 2004. The full dataset is divided into 21 groups, each containing 10-30 males and a comparable number of females (though the numbers are not necessarily identical). Within any group, each agent engaged in a four-minute speed dating with every agent of the opposite sex in the same group. After the dates, participants were asked to rate their partners on six attributes: attractiveness, sincerity, intelligence, fun, ambition, and shared interests, as well as to report the importance they assigned to each attribute. Each participant also provided overall scores of their partners on a 1–10 scale. The dataset also includes self-reported questionnaires and background information collected from participants.

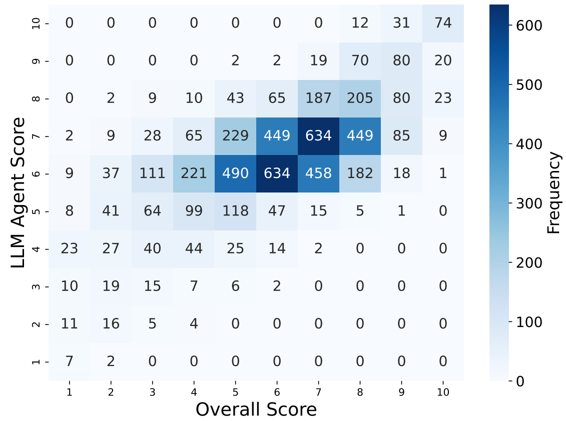
**Target LLMs.** For quantal response validation, we fix a single LLM model: GPT-4-Turbo (Achiam et al., 2023). For all matching and mechanism experiments, we consider 5 representative LLMs: Claude-Sonnet-4 (Bai et al., 2022), GPT-4-Turbo, Gemini-2.5-Flash (Team et al., 2023), Qwen2.5-72B-Instruct (Yang et al., 2025a), and DeepSeek-R1 (Guo et al., 2025).

**Matching Methods.** We compare three matching mechanisms in our experiments: a classical Gale–Shapley matching based on the deterministic preference rankings induced by the overall scores in the dataset, a Hybrid GS-LLM Matching method, and a random matching baseline.

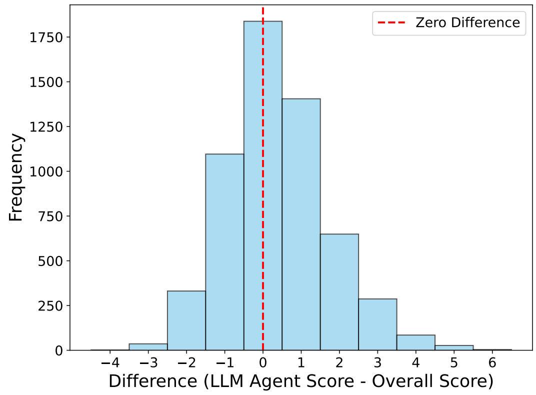
**Prompting Protocol.** To elicit switching behavior from LLMs in a controlled manner, we design two task-oriented prompts corresponding to two different states of an agent. When an agent is in a "single" state, he or she decides whether to initiate a relationship with a potential partner. When an agent is in an "engaged" state, he or she decides whether to leave the current partner in favor of a new one. All prompts use structured inputs and enforce strictly formatted binary outputs, which helps reduce hallucination and ensures comparability. Full prompt templates are provided in Appendix A.

### 4.2 Experimental Analysis of Quantal Response Dynamics

Given the six attribute ratings of a dating pair, we prompt the LLM models to predict the participants'




(a) Joint distribution of LLM and human scores.




(b) Distribution of score differences.

Figure 3: Comparison between LLM-generated scores and human overall evaluations.

evaluations of their dating partners from 1 to 10 for their partners. We compare the LLM-generated scores with the original overall scores. As shown in Figure 3, the scores produced by the LLM agents closely match the ground-truth evaluations. In particular, the majority of score differences are no larger than 1. This consistency is further reflected in the Mean Absolute Error (MAE) between the two sets of scores, which is 1.03, close to 1, indicating that the LLM predictions align well with human-reported overall scores. Full results are provided in E.1. Across models, LLM-generated scores exhibit moderate but consistent correlations with human responses. Overall, this sanity check suggests that LLMs internalize patterns that are similar to human participants.

### 4.3 Probabilistic Switching Model

**Fitting the quantal response model.** We first examine whether agents' switching behavior can be well described by a quantal response model. The decision setting naturally aligns with the logit quantal response, in which the probability of choosing an option is a smooth function of the underlying utility difference and the choice is binary. This formulation is formally equivalent to standard logistic regression.

Table 1: Estimated logit-QR parameters for different LLM agents (from Fig. 4).

| LLM Model | $\lambda$ | $\beta$ | $-\beta/\lambda$ |
|---|---|---|---|
| GPT-4-Turbo | 0.09 | -2.48 | 27.56 |
| Claude-Sonnet | 0.07 | -1.51 | 21.57 |
| DeepSeek-R1 | 0.13 | -0.24 | 1.85 |
| Gemini-Flash | 0.10 | -0.05 | 0.50 |
| Qwen-72B | 0.12 | -1.38 | 11.50 |

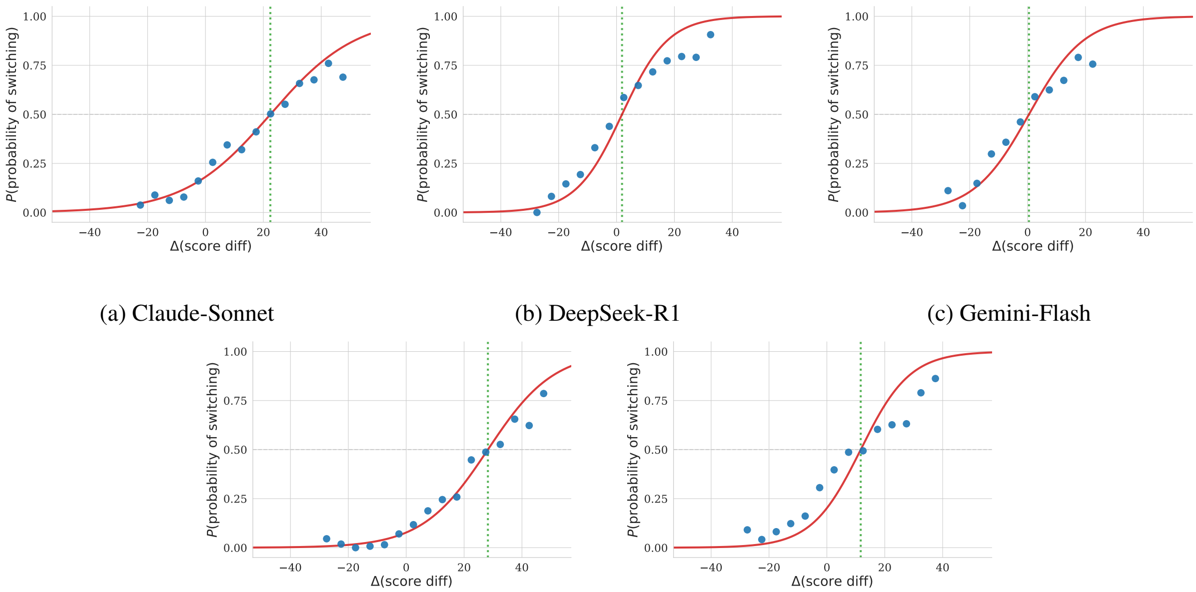
Formally, let the score difference for agent $i$ under LLM model $a$ be $\Delta_i^a(k) = s_i^a(k) - s_i^a(j)$ where $j$ is $i$'s current partner and $k$ is the new proposer. We model the switching probability as

$$p^a(\Delta) = \frac{1}{1 + e^{-\lambda^a \cdot \Delta + \beta^a}},$$

where $\lambda^a$ controls the sensitivity to utility differences, and $\beta^a$ shifts the curve. Given binary observations $y \in \{0, 1\}$, we estimate $(\lambda^a, \beta^a)$ by minimizing the cross-entropy loss:

$$\mathcal{L}(\lambda^a, \beta^a) = -\sum_n \Big[ y_n \log p^a(\Delta_n) + (1 - y_n) \log \big(1 - p^a(\Delta_n)\big) \Big].$$

Across all LLMs, the observed relationship between switching probability and score difference exhibits a smooth S-shaped curve, closely matching the form predicted by quantal response models. The results can be found in Figure 4. This result suggests that LLM-induced decision behavior naturally conforms to response patterns with bounded rationality, providing empirical support for our probabilistic switching model. Beyond that, Table 1 also reveals different behavioral patterns across LLM agents.The sensitivity parameter $\lambda^a$ governs how sharply switching probabilities respond to score differences, while the offset $\beta^a$ captures inertia toward maintaining the current match. Models with more negative $\beta^a$ exhibit stronger status-quo bias, requiring a larger score advantage before switching becomes likely. As summarized, GPT-4-Turbo and Claude-Sonnet show stronger sensitivity than Gemini-Flash, while Deepseek-R1 exhibits a comparatively different offset, indicating a different baseline tendency before utility difference takes effect. In particular, the ratio $-\beta^a/\lambda^a$




(a) Claude-Sonnet (b) DeepSeek-R1 (c) Gemini-Flash

(d) GPT-4-Turbo (e) Qwen2.5-72B

Figure 4: Empirical switching probabilities and fitted quantal response curves for five large language models. Each sub-figure plots the empirical probability of switching. The vertical dashed line indicates the indifference point ($p = 0.5$), corresponding to the estimated utility of the single state.

corresponds to the threshold implied by the quantal response model, i.e., the score difference at which switching or not are equally likely. This threshold may reflect how large a relative utility advantage is required before an agent becomes inclined to switch. As shown, GPT-4-Turbo and Claude-Sonnet exhibit larger thresholds, indicating that these models tend to favor maintaining the current match. In contrast, Gemini-Flash and DeepSeek-R1 have much smaller thresholds, suggesting a greater willingness to switch when even modest improvements are available.

**Estimating the single state utility.** While the score of each candidate partner is available, the utility associated with remaining single is unobserved. We associate it with a model-specific score $s_i^a(i)$, following the standard convention that an unmatched agent is matched to itself.

When agent $i$ considers a proposer $j$, the acceptance probability under LLM model $a$ is

$$\Pr(j) = \frac{1}{1 + e^{-\left(\lambda^a (s_i^a(j) - s_i^a(i)) + \beta^a\right)}},$$

where $(\lambda^a, \beta^a)$ are fixed from the previously estimated QR model, and $s_i^a(j)$ is known. We estimate the unknown $s_i^a(i)$ by minimizing the cross-entropy loss with respect to $s_i^a(i)$:

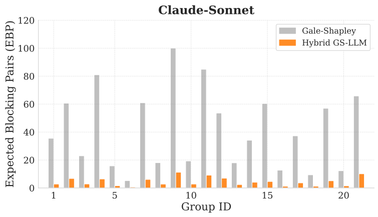
$$\mathcal{L}(s) = \arg\min_s \sum_n \Big[ y_n \log p_{n,i}^a(s) + (1 - y_n) \log\left(1 - p_{n,i}^a(s)\right) \Big],$$

where $y_n$ denotes the observed decision and $p_{n,i}^a(s)$ is the corresponding QR probability with $s = s_i^a(i)$.

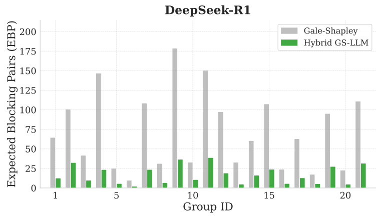
Table 2: Estimated single-state scores across LLM agents (rescaled to a 10-point scale).

| LLM Agent | score of single state |
|---|---|
| GPT-4-Turbo | 1.09 |
| Claude-Sonnet | 1.36 |
| Qwen-72B | 2.38 |
| DeepSeek-R1 | 2.66 |
| Gemini-Flash | 3.61 |

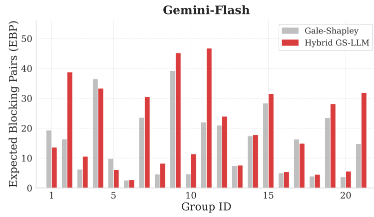
Estimated single-state scores for different LLM agents are summarized in 2. The scores reveal substantial variation in baseline selectiveness, reflecting the implicit utility an agent assigns to remaining unmatched. Models such as GPT-4-Turbo and Claude-Sonnet strongly favor being matched. In practice, these models tend to accept relationships even when the perceived match quality is modest. In contrast, Gemini-Flash exhibits a higher



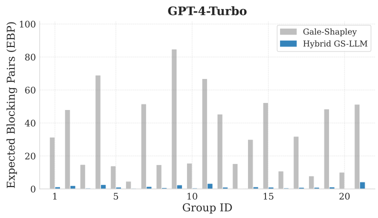

(a) Claude-Sonnet



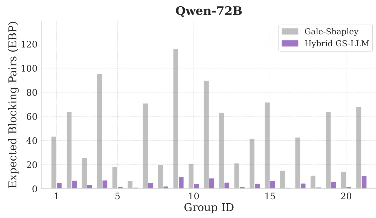

(b) DeepSeek-R1




(c) Gemini-Flash




(d) GPT-4-Turbo




(e) Qwen2.5-72B

Figure 5: Comparison of expected blocking pairs (EBP) between Gale–Shapley and the Hybrid GS–LLM matching across 21 speed-dating groups. Each pair of points corresponds to one group, with the two methods evaluated under identical score constructions. While the Gale–Shapley algorithm guarantees deterministic stability, the Hybrid GS–LLM method achieves lower EBP in most groups, indicating improved stability under probabilistic switching behavior.

tolerance for remaining single. Qwen-72B and DeepSeek-R1 lie between these extremes.

Together with the estimated quantal response parameters, these results highlight systematic differences in how LLM agents balance commitment and selectiveness in two-sided matching environments.

### 4.4 Expected Blocking Pairs under Different Matching Methods

For each matching instance, EBP is computed using switching probabilities estimated from the behavioral model, and to account for stochasticity, we run each matching method across multiple instances, and examine the best-performing outcomes. We find that the hybrid GS-LLM method achieves lower EBP than the classical Gale-Shapley method in a substantial fraction of classes, indicating systematic rather than incidental improvements. Figure 5 directly compares the hybrid method and the classical Gale-Shapley algorithm under identical score constructions. The Gale-Shapley algorithm does not account for probabilistic deviations and therefore can incur relatively lower stability once agents behave stochastically. These results suggest that adapting the matching mechanism to stochastic behavior yields matchings that are more robust in expectation.

Beyond aggregate improvements, we can also see in Figure 5 that the reduction in EBP is more pronounced in settings where agent behavior is more stochastic, such as the model GPT-4-Turbo and Claude-Sonnet. When agents exhibit near-deterministic responses, i.e., with a relatively greater $\lambda$, classical Gale-Shapley already produces out comes that are stable even in expectation.

As an additional reference point, we also compare all methods against a random matching baseline that ignores preference information entirely. These results confirm that the observed improvements of the Hybrid GS–LLM method are not driven by trivial randomness. For brevity, we report the detailed comparisons with random matchings in Appendix E.2.

## 5 Conclusion

We find that LLMs exhibit structured decision behavior in two-sided matching settings. Also, LLMs exhibit partner-choice patterns that align closely with bounded rationality as characterized by quantal response models. When decisions are stochastic, classical deterministic stability fails to capture realized outcomes, while matchings produced by LLM agents that account for probabilistic behavior achieve lower expected instability. These findings suggest that LLMs serve as a meaningful tool for studying preference uncertainty and stability in matching markets.

## Limitations

Our work has several limitations that point to directions for future research.

- We may not capture richer interpersonal dynamics such as context-dependent attraction or long-term compatibility, because the preference scores used in our framework are constructed by aggregating a fixed set of observable attributes.
- We do not extend our analysis beyond static one-to-one matching markets, and leave many-to-one settings, dynamic participation, and strategic manipulation as directions for future work.
- We do not model strategic dependencies or social correlations between agents' switching decisions, instead assuming independence to obtain a tractable and interpretable measure of probabilistic stability.

## References


Alberto Acerbi and Joseph M Stubbersfield. 2023. Large language models show human-like content biases in transmission chain experiments. *Proceedings of the National Academy of Sciences*, 120(44):e2313790120.

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, and 1 others. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Ahmet Alkan and Kemal Yildiz. 2023. Equitable stable matchings under modular assessment. In *Proceedings of the 24th ACM Conference on Economics and Computation*, EC '23, page 64, New York, NY, USA. Association for Computing Machinery.

Itai Ashlagi, Jiale Chen, Mohammad Roghani, and Amin Saberi. 2025. Stable matching with interviews. In *ITCS*.

Haris Aziz, Péter Biró, Serge Gaspers, Ronald de Haan, Nicholas Mattei, and Baharak Rastegari. 2020. Stable matching with uncertain linear preferences. *Algorithmica*, 82(5):1410–1433.

Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones, Anna Chen, Anna Goldie, Azalia Mirhoseini, Cameron McKinnon, and 1 others. 2022. Constitutional ai: Harmlessness from ai feedback. *arXiv preprint arXiv:2212.08073*.

Matthias Bentert, Niclas Boehmer, Klaus Heeger, and Tomohiro Koana. 2023. Stable matching with multi-layer approval preferences: approvals can be harder than strict preferences. *Games and Economic Behavior*, 142:508–526.

Valerio Capraro, Roberto Di Paolo, and Veronica Pizziol. 2025. A publicly available benchmark for assessing large language models' ability to predict how humans balance self-interest and the interest of others. *Scientific Reports*, 15(1):21428.

Vanessa Cheung, Maximilian Maier, and Falk Lieder. 2025. Large language models show amplified cognitive biases in moral decision-making. *Proceedings of the National Academy of Sciences*, 122(25):e2412015122.

Changde Du, Kaicheng Fu, Bincheng Wen, Yi Sun, Jie Peng, Wei Wei, Ying Gao, Shengpei Wang, Chuncheng Zhang, Jinpeng Li, and 1 others. 2025. Human-like object concept representations emerge naturally in multimodal large language models. *Nature Machine Intelligence*, pages 1–16.

Antonino Ferraro, Antonio Galli, Valerio La Gatta, Marco Postiglione, Gian Marco Orlando, Diego Russo, Giuseppe Riccio, Antonio Romano, and Vincenzo Moscato. 2024. Agent-based modelling meets generative ai in social network simulations. In *International Conference on Advances in Social Networks Analysis and Mining*, pages 155–170. Springer.

Raymond Fisman, Sheena S Iyengar, Emir Kamenica, and Itamar Simonson. 2006. Gender differences in mate selection: Evidence from a speed dating experiment. *The Quarterly Journal of Economics*, 121(2):673–697.

David Gale and Lloyd S Shapley. 1962. College admissions and the stability of marriage. *The American mathematical monthly*, 69(1):9–15.

Yitian Gao, Jiaxue Li, Junjie Luo, and Yiheng Zhang. 2025. Minimizing blocking agents for stable matching with partial approval information. In *International Workshop on Frontiers in Algorithmics*, pages 293–306. Springer.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, and 1 others. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

John J Horton. 2023. Large language models as simulated economic agents: What can we learn from homo silicus? Technical report, National Bureau of Economic Research.

Jiaxin Liu, Yixuan Tang, Yi Yang, and Kar Yan Tam. 2025. Evaluating and aligning human economic risk preferences in llms. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 18185–18199.

Nunzio Lorè and Babak Heydari. 2024. Strategic behavior of large language models and the role of game structure versus contextual framing. *Scientific Reports*, 14(1):18490.

Richard D McKelvey and Thomas R Palfrey. 1995. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38.

Trinh Bao Ngoc, Dang Nhat Quang, Dang Tien Dat, Le Anh Phan, Nguyen Ngoc Phuong Khanh, Ha Thi Thanh Thao, and Nguyen Xuan Thang. 2024. Combination of gale-shapley and pesa-ii algorithm in student-university match. In *Conference on Information Technology and its Applications*, pages 147–159. Springer.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pages 1–22.

Boris Pittel. 2025. On constrained matchings, stable under random preferences. *Random Structures & Algorithms*, 67(4):e70036.

Alvin E Roth and Marilda Sotomayor. 1992. Two-sided matching. *Handbook of game theory with economic applications*, 1:485–541.

James WA Strachan, Dalila Albergo, Giulia Borghini, Oriana Pansardi, Eugenio Scaliti, Saurabh Gupta, Krati Saxena, Alessandro Rufo, Stefano Panzeri, Guido Manzi, and 1 others. 2024. Testing theory of mind in large language models and humans. *Nature Human Behaviour*, 8(7):1285–1295.

Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, Katie Millican, and 1 others. 2023. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*.

Yinghui Wen, Aizhong Zhou, and Jiong Guo. 2022. Position-based matching with multi-modal preferences. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, pages 1373–1381.

Yutong Xie, Qiaozhu Mei, Walter Yuan, and Matthew O Jackson. 2025. Using large language models to categorize strategic situations and decipher motivations behind human behaviors. *Proceedings of the National Academy of Sciences*, 122(35):e2512075122.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, and 1 others. 2025a. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

Yang Yang, Boon Han Lim, Yicheng Xu, and Yong Zhang. 2025b. Spectral clustering and multi-modal stable matching for enhanced management of photovoltaic stations and virtual power plants. In *2025 9th International Conference on Green Energy and Applications (ICGEA)*, pages 1–5. IEEE.


# A Prompt Templates

### A.0.1 LLM-Generated Preference Prompts

We elicit LLM-implied preference scores based on the six attributes from the Speed-Dating dataset. Rather than using numerical ratings, we represent both the partner's attribute evaluations and the agent's importance preferences using qualitative textual descriptions (e.g., "slightly satisfied", "satisfied"). The agent is instructed to synthesize these textual descriptions and output exactly one integer between 1 and 10 as the overall preference score. Identical system messages and decoding settings are applied across all LLMs.

**LLM-Generated Preference Prompts**

```
Assume you went on a speed date
with someone you have not
previously met. You are a
{}-year-old {}, and your date is {}
years old. Your field of study is
{}.

After the date, you evaluated {}
on six dimensions. These are
attractiveness: {}, sincerity: {},
intelligence: {}, being funny: {},
ambition: {}, shared interests:{}.

Now please rate him with a score
from 1 to 10 and answer me (I
would rate my partner)+(your
score). You must follow this rule.

And would you like to have one more
date with him? Answer (Yes) or
(No). You must follow this rule.
```

### A.0.2 Switching Prompts

To elicit partner-switching behavior, we present LLM agents with controlled comparison prompts in which they decide whether to remain with a current partner or switch to a new proposer. Each prompt specifies the six attribute ratings of the current partner and those of the new proposer, again on the original 1–10 scale. The agent outputs a binary decision indicating whether it would switch. Across all switching prompts, the textual template is fixed and only the attribute values vary. For analysis, we summarize each scenario by the overall score difference

**Switching Prompts**

Assume you are a {age}-year-old woman with career type {career}, currently matched with a partner.

Current Partner Info: Your current partner's scores (out of 10):

- attractiveness: {current_partner_score},
- sincerity: {current_partner_score},
- intelligence: {current_partner_score},
- being funny: {current_partner_score},
- ambition: {current_partner_score},
- shared interests: {current_partner_score}.

New Proposer Info: A new man is courting you with scores:

- attractiveness: {proposer_score},
- sincerity: {proposer_score},
- intelligence: {proposer_score},
- being funny: {proposer_score},
- ambition: {proposer_score},
- shared interests: {proposer_score}.

Your Preferences: Your importance weights are:

- attractiveness: {importance},
- sincerity: {importance},
- intelligence: {importance},
- being funny: {importance},
- ambition: {importance},
- shared interests: {importance}.

**Decision:** Would you leave your current partner for the new man? Your answer must be (YES) or (NO) followed by your reason. You must FOLLOW THIS RULE! You MUST NOT SAY other words.

### A.0.3 Single-State Prompts

We treat being unmatched (single) as a valid outside option and elicit acceptance decisions relative to this option using single-state prompts. Given the six attribute ratings of a proposer, the agent decides whether to accept the proposer or remain single, producing a binary output. The implied utility of the single state is estimated by identifying the proposer score at which the acceptance probability is closest to 0.5. This procedure serves only as an empirical calibration of the outside option and does not impose any deterministic acceptance thresholds during matching.

**Single-State Prompts**

Assume you are a {age}-year-old single man, and there is a {age_o}-year-old woman who is courting you. Your career type is {career}. You evaluated her on six dimensions with scores out of 10. These are:

- attractiveness: {score}
- sincerity: {score}
- intelligence: {score}
- being funny: {score}
- ambition: {score}
- shared interests: {score}

Your importance weights for these attributes are:

- attractiveness: {importance},

- sincerity: {importance},
- intelligence: {importance},
- being funny: {importance},
- ambition: {importance},
- shared interests: {importance}.

**Task:** Would you be willing to date this woman? And also tell me the reason.

**Constraint:** Your answer must be (YES) or (NO) followed by your reason. You must FOLLOW THIS RULE! You MUST NOT SAY other words.

## B Dataset and Score Construction

We use the Speed-Dating dataset, which consists of 21 speed-dating events with uneven numbers of male and female participants. Each participant rates every opposite-sex partner on six attributes using a 1–10 scale. We remove samples with missing attribute values and do not apply any normalization or rescaling to the remaining ratings. For each agent–partner pair, we obtain an overall preference score by directly querying the LLM as described in Appendix A. Each agent's ordinal preference list is then constructed by sorting potential partners according to these integer scores, with ties broken deterministically by participant ID. Because unmatched outcomes are allowed, agents on the proposing side sequentially propose to all candidates in their preference list, and any remaining agents are assigned to the `single` state.

## C Quantal Response Estimation Details

We summarize LLM-induced stochastic switching behavior using a logit quantal response specification. For each switching scenario, we define the utility difference $\Delta = s_{\text{new}} - s_{\text{cur}}$ based on the integer preference scores generated by the LLM. The probability of switching is modeled as

$$p(\text{switch} \mid \Delta) = \frac{1}{1 + \exp(-(\lambda\Delta + \beta))},$$

where $\lambda$ controls sensitivity to score differences and $\beta$ captures a baseline offset. Parameters $(\lambda, \beta)$ are estimated by minimizing the logistic loss over all observed switching decisions. For stability evaluation, we assume independence between the switching decisions of the two agents in a potential pair; under this assumption, the probability of a blocking deviation is given by the product of the two agents' switching probabilities.

## D Hybrid GS–LLM Method Implementation Details

The Hybrid GS–LLM procedure follows the proposal-based structure of the Gale–Shapley algorithm with deterministic proposal order determined by agent IDs. Proposers sequentially propose to candidates according to their LLM-induced preference rankings. Upon receiving a proposal, a receiver samples an accept or reject decision according to its quantal-response-based switching probability relative to its current assignment or the `single` option. Each proposer contacts each receiver at most once, ensuring termination after a finite number of proposals. Because unmatched outcomes are permitted, agents who are never accepted remain in the `single` state, which is treated as a valid assignment when computing stability metrics.

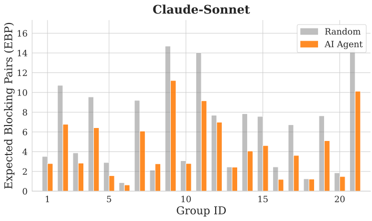
## E Additional Results

### E.1 Sanity Check

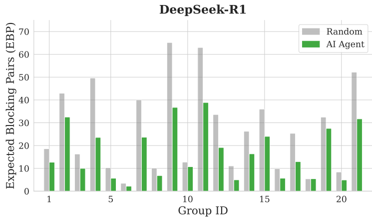
Table 3: Sanity check on preference reconstruction.

| Metric | GPT-based Agent |
|---|---|
| MAE ↓ | 1.03 |
| Pearson Correlation | 0.663 |
| Spearman Correlation | 0.609 |

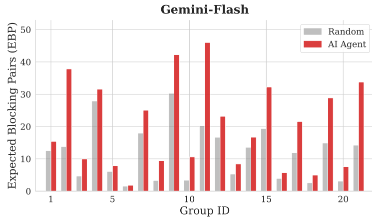
Given the six attribute ratings of each dating pair, we prompt the LLM agent to output an overall preference score and compare it with the observed human "like" responses. Table 3 reports the mean absolute error (MAE) and rank-based correlations between LLM-generated scores and human judgments. The results show moderate but consistent correlations, indicating that LLM-implied preferences are meaningfully aligned with human evaluations while remaining noisy and imperfect. This level of alignment supports the use of LLMs as stochastic preference agents, rather than deterministic predictors of individual choices.



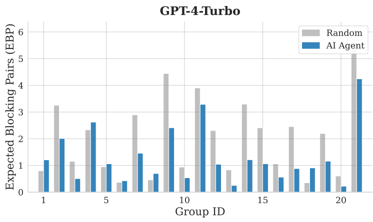

(a) Claude-Sonnet



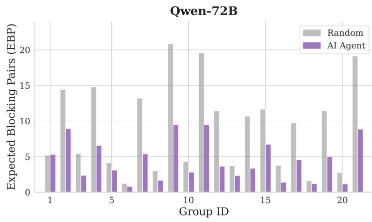

(b) DeepSeek-R1




(c) Gemini-Flash




(d) GPT-4-Turbo




(e) Qwen2.5-72B

Figure 6: Expected blocking pairs (EBP) under random matching and LLM-based matching across 21 speed-dating groups. Lower values indicate more stable outcomes. For the majority of groups, matchings induced by LLM-based decision models exhibit substantially lower EBP than random matching, suggesting that LLM-generated preferences capture non-random, structured decision behavior that improves probabilistic stability.

## E.2 Random Baseline

As a baseline, we consider random matchings that ignore all preference information. For each event, we generate random matchings by pairing agents uniformly at random while allowing unmatched agents to remain single when group sizes differ. For each random matching, we compute the Expected Blocking Pairs (EBP) metric using the same quantal-response-based switching probabilities as in the main experiments. We repeat this process multiple times per event and report the resulting EBP distributions. The results are shown in Figure 6. As expected, random matchings exhibit substantially higher expected instability, reflecting the absence of structured preference alignment.