Seeing Who Liked You: Information Design, Rating Risk, and Willingness to Pay in Dating Apps

Your Name

November 14, 2025

1 Theoretical Framework

In this section we formalize a simple decision problem that links users' beliefs, effort choices, and the observable OkCupid profiles in our cross-sectional data.

1.1 Users, Ratings, and Goals

Each user i is described by two components:

- a continuous *rating* $r_i \in [0, 1]$ capturing how desirable they are as a partner;
- a *relationship goal* $g_i \in \{S, L\}$, where S denotes short-term orientation and L denotes long-term orientation.

In the empirical part, r_i will be constructed as a scalar index from observable profile features (demographics, education, lifestyle variables, and possibly profile completeness) and normalized to $[0, 1]$. The goal g_i will be proxied from text and profile fields (e.g. LTR-like vs. casual-like language).

User i has an aspiration level $a_i \geq 0$ and ideally wants to match with partners whose rating is at least

$$r_j \geq r_i + a_i$$

and who share her goal ($g_j = g_i$). Let $v(r_j)$ denote the payoff from a goal-matched relationship with partner j. We assume $v(\cdot)$ is increasing and continuous in r_j (for example, $v(r_j) = v_0 + v_1 r_j$ with $v_1 > 0$), so that higher-rated partners are more attractive, all else equal.

1.2 Market Composition and Goal Beliefs

Users are uncertain about the share of long-term oriented partners in their local dating market. For a given market m (e.g. city $\times$ orientation), let

$$\rho_m \equiv \Pr(g_j = L \mid j \text{ in market } m)$$

denote the (unknown) share of long-term oriented users in that market.

A newcomer in market m arrives with a prior belief

$$\rho_m \sim \text{Beta}(\alpha_0, \beta_0),$$

so that the prior mean share of long-term users is

$$\mathbb{E}[\rho_m] = \frac{\alpha_0}{\alpha_0 + \beta_0}.$$

When the user browses the app, she sees a sequence of profiles $t = 1, 2, \dots$ with observed goal labels $\tilde{g}_t \in \{L, S\}$ (for example, text suggesting long-term vs. casual intent). After observing k_L profiles labeled L and k_S labeled S, her posterior mean becomes

$$\widehat{\rho}_m = \mathbb{E}[\rho_m \mid k_L, k_S] = \frac{\alpha_0 + k_L}{\alpha_0 + \beta_0 + k_L + k_S}. \tag{1}$$

The platform's recommendation algorithm determines which profiles are shown. If newcomers are disproportionately shown high-rated users who appear short-term oriented, then k_S will be large relative to k_L, and the posterior $\widehat{\rho}_m$ in (1) becomes pessimistic, even in markets where the true share of long-term users is substantial.

1.3 Learning About Acceptance When Aiming Upwards

Users are also uncertain about their own probability of success when reaching out to more desirable partners. Fix an aspiration gap $\Delta > 0$ and consider outreach attempts to partners with

$$r_j \geq r_i + \Delta \quad \text{and} \quad g_j = L.$$

Let p_Δ denote the probability that such an attempt leads to reciprocation (a match).

User i starts with a prior belief

$$p_\Delta \sim \text{Beta}(a_0, b_0),$$

with prior mean $\mathbb{E}[p_\Delta] = a_0/(a_0 + b_0)$.

Each time she likes or messages a partner in this aspirational region, she observes $y_t \in \{0, 1\}$, where $y_t = 1$ indicates reciprocation and $y_t = 0$ indicates rejection or nonresponse. After k_Δ successes and m_Δ failures, her posterior mean success probability is

$$\widehat{p}_\Delta = \mathbb{E}[p_\Delta \mid k_\Delta, m_\Delta] = \frac{a_0 + k_\Delta}{a_0 + b_0 + k_\Delta + m_\Delta}. \tag{2}$$

If the platform initially shows many high-r profiles to newcomers, users are likely to accumulate failures m_Δ early on. This drives $\widehat{p}_\Delta$ down and can make them quickly pessimistic about their chances when aiming upwards.

1.4 Effort Choice and Revealed Preferences

We interpret each user's observable profile as the result of an effort choice taken after some experience in the app. For a long-term oriented user ($g_i = L$), we consider two stylized strategies:

1. **Goal-consistent, aspirational strategy** $A = I$: the user mainly targets partners with $g_j = L$ and $r_j \geq r_i + \Delta$ ("aiming upwards") and chooses a high profile effort level e_I (longer essays, more complete information, clear mention of long-term goals). Let $c(e_I) > 0$ be the cost of this effort.

2. **Adapted, low-effort strategy** $A = C$: the user broadens or shifts targeting to lower-r_j or more casual-looking partners and chooses a low effort level e_C (short text, ambiguous or casual goals, many missing fields), with $c(e_C) < c(e_I)$.

Let $\psi(\cdot)$ denote the psychological cost of rejection, increasing in the expected rejection probability. Under $A = I$, the expected utility from one outreach attempt is

$$U_I = \underbrace{\mathbb{E}\big[v(r_j) \mid r_j \geq r_i + \Delta,\, g_j = L\big] \cdot \widehat{p}_\Delta}_{\text{expected match value}} - c(e_I) \ - \ \psi\big(1 - \widehat{p}_\Delta\big). \tag{3}$$

Under $A = C$, the user mainly targets partners with higher acceptance probability p_C (e.g. smaller rating gap, more casual goals). Let $\bar{v}_C$ be the expected value of such matches. Then

$$U_C = \bar{v}_C\, p_C \ - \ c(e_C) \ - \ \psi\big(1 - p_C\big). \tag{4}$$

The chosen strategy (revealed preference) is

$$A^\star = \begin{cases} I & \text{if } U_I \geq U_C, \\ C & \text{if } U_I < U_C. \end{cases}$$

By construction, a long-term oriented user with an aspiration gap Δ intrinsically prefers the matches generated by $A = I$: $\mathbb{E}[v(r_j) \mid r_j \geq r_i + \Delta, g_j = L] > \bar{v}_C$. However, when early experiences in the app lead to pessimistic beliefs

$$\widehat{\rho}_m \text{ low (few observed long-term profiles)} \quad \text{and} \quad \widehat{p}_\Delta \text{ low (many failed attempts)},$$

the rejection cost $\psi(1 - \widehat{p}_\Delta)$ becomes large and the match-value term in (3) becomes small. In this region it is possible that

$$\bar{v}_C < \mathbb{E}[v(r_j) \mid r_j \geq r_i + \Delta, g_j = L] \quad \text{but} \quad U_I < U_C,$$

so that the user optimally switches to $A^\star = C$ despite truly preferring long-term, higher-rated partners.

In cross-sectional data, this appears as a wedge between *true* preferences and *revealed* behavior: long-term oriented users may display low-effort, ambiguous, or casual-looking profiles because their beliefs have become pessimistic.

1.5 Mapping to the OkCupid Data

Our dataset is cross-sectional and does not contain the history of swipes or messages. We therefore interpret each observed profile as the outcome of the effort choice above, after some (unobserved) sequence of belief updates. We approximate the theoretical objects using proxies:

- **Effort E_i:** essay word count, number of non-empty essays, and profile completeness summarized into an effort index.
- **Goal type g_i:** text-based classification of profiles into LTR-like, casual-like, or ambiguous.
- **Rating r_i:** a continuous index constructed from observable characteristics (demographics, education, lifestyle, and possibly effort), normalized to $[0, 1]$.
- **Market beliefs $(\widehat{\rho}_m, \widehat{p}_\Delta)$:** for each market m we compute the observed share of LTR-like profiles among potential partners, the distribution of their ratings, and the share of LTR-like profiles among high-rated partners. These market-level indices proxy for the environments that are more likely to generate pessimistic beliefs and thus lower effort.

In the empirical specifications below, we test whether users in markets with a lower share of long-term profiles and more intense competition exhibit lower effort and more "adapted" (casual-like) self-presentation, consistent with the comparative statics implied by the model.