A PRE-NORMALIZATION

Though we focus on the similarity between q_t and d, magnitudes of q_t do matter in many practical situations. For example, if the decision-maker prefers students with balanced math and English skills and there are two "balanced" students, the decision-maker will certainly prefer the one with higher scores. Therefore, we propose a **pre-normalization procedure** to incorporate magnitude into account. Specifically, we add an additional dimension representing the unobservable "irrelevant attributes" to q_0 and obtain a $m+1$ dimensional complete qualification profile. Meanwhile, we add an additional dimension to the ideal qualification profile d with 0 as its value; the new ideal profile becomes $[d; 0]$. Then we can make the following natural assumption:

Assumption A.1. After adding the dimension of "irrelevant attribute", for all agents, the norms of their complete qualification profiles are the same.

Assumption A.1 has been supported by literature in machine learning Liu et al. (2022) and social science Holmstrom & Milgrom (1991). The "irrelevant" dimension demonstrates all other skills that belong to an agent but are not important to the decision. Therefore, competency in relevant/measurable attributes implies weakness in irrelevant/immeasurable attributes and the length of the complete qualification profile stays the same for all agents. With Assumption A.1 and the distribution of q_0 as Q, we formalize the **pre-normalization** procedure in Algorithm 1.

Algorithm 1 Pre-normalization procedure

Require: Joint distribution Q for q_0, n agents with $\{q_0^i\}_{i=1}^{i=n}$ where $q_0^i \in [0,1]^m$, $d \in [0,1]^m$.
Ensure: Normalized $\{q_0^i\}_{i=1}^{i=n}$ (i.e., $q_0^i \in [0,1]^m$ and $\|q_0^i\| = 1$), new $d \in [0,1]^{m+1}$.

1: $d = [d, 0]$.
2: According to Q, find the largest norm $K = \max_{q_0 \sim Q} \|q_0^i\|$ of original profiles.
3: **for** $i \in \{1, \dots n\}$ **do**
4: Calculate norm difference $z^i = \sqrt{K^2 - \|q_0^i\|_2^2}$.
5: $q_0^i = \frac{[q_0^i; z^i]^T}{K} \in [0,1]^{m+1}$.
6: **end for**

B DISCUSSION AND GENERALIZATION OF EQUATION 1

Motivating examples of the dynamics in equation 1. In the main paper, we assume the influence of the one-time effort k is persistent and will enable q_t changes gradually during each round. This is well-supported by the following examples:

1. *Creditworthiness*: To improve creditworthiness, an individual may learn that an ideal profile would be a person with a constant high income and long-lasting good credit history. Therefore, she may exert a significant effort to find a job with a high salary. However, the effort will take several months or even one year for her to finally build up the ideal profile because she needs to work for a while to receive money and build a competitive credit history.
2. *Job application*: An individual who wants to apply for a technology company may learn about the skill set of an ideal candidate from several resources (e.g., the job description, alumni who work at the company, info session) and then exert a significant effort to study the required knowledge. However, it still takes time for her to do exercises and master the skills, resulting in a delay of finally being qualified.

Model generalization when k changes with t. In the main paper, k_t is always equal to k, demonstrating the effort has a consistent and persistent effect on the improvement of an individual. According to Lemma 2.1, the similarity x_t approaches 1 at an exponential rate. Thus, the case of $k_t > k$ is not interesting since the convergence is faster and it may not make sense in practice that the effort can be increasingly effective as time goes on. However, in reality, it may be possible that k_t is decreasing. This is a "middle-point" case between the regular improvement in equation 1 and the forgetting mechanism equation 9, which may illustrate the "tiredness" when agents stick to improve. However, we can prove that when k_t decreases linearly (i.e., $k_t = \Theta(\frac{k}{t})$), the similarity x_t can only converge to 1 at a speed $\Theta(t^k)$.

Theorem B.1. *When k_t decreases linearly (i.e., $k_t = \Theta(\frac{k}{t+1})$), x_t converges to 1 at a rate $\Theta(t^k)$*

We prove Thm. B.1 in App. H.5. Basically, this result illustrates that the agents will still improve to be qualified if k_t decreases at a linear rate. Specifically, we can rewrite the equation 2 as:

$$x_t^{-2} - 1 = \frac{(x_0)^{-2} - 1}{(t+1)^{2k}} \tag{12}$$

From equation 12, we can derive similar results of the agents' best responses and work out the thresholds for them to improve.

C Related Work

C.1 Strategic Manipulation

Though our work primarily lies in proposing a new model for improvement behaviors, the problem settings are also closely related to strategic classification problems Hardt et al. (2016a); Ben-Porat & Tennenholtz (2017); Dong et al. (2018); Braverman & Garg (2020); Sundaram et al. (2021); Jagadeesan et al. (2021); Ahmadi et al. (2021); Eilat et al. (2022); Horowitz & Rosenfeld (2023). Hardt et al. (2016a) formulated classification problems with strategic manipulation as a Stackelberg game with deterministic cost functions, where the decision maker optimizes classification accuracy based on individuals' best responses. Afterwards, more sophisticated analytical frameworks were proposed Dong et al. (2018); Braverman & Garg (2020); Jagadeesan et al. (2021). Dong et al. (2018) proposed an online algorithm for strategic classification, and Braverman & Garg (2020) added randomness to strategic classifiers. On the other hand, Sundaram et al. (2021) analyzes the statistical learnability of strategic classification with an SVC classifier. Jagadeesan et al. (2021) relaxed the *standard microfoundations* assumption where individuals are perfectly rational to *alternative microfoundations* where a proportion of individuals may not be strategic, and proposed a *noisy response model* to tackle the new problem. Zhang et al. (2022) studied the setting where the decision maker and individuals only have knowledge of the feature distributions as random variables. Thus, the strategic manipulation corresponds to a distribution shift and its cost is also a random variable. Eilat et al. (2022) considered the setting where individual responses are dependent and the classifier is learned through *graph neural networks*.

C.2 Improvement

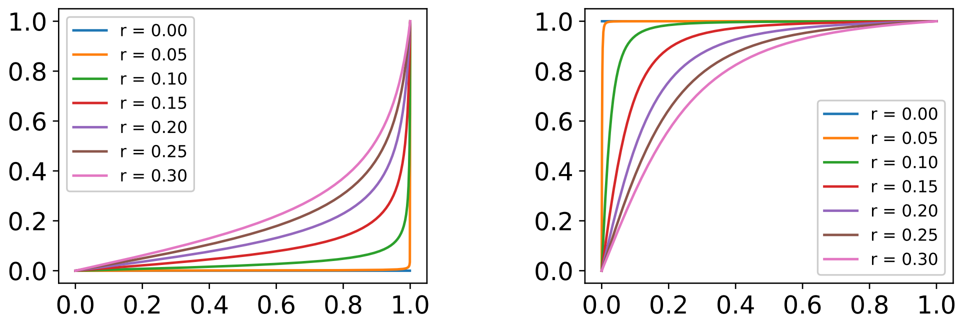
However, there are other literature considering improvement behaviorLiu et al. (2019); Rosenfeld et al. (2020); Shavit et al. (2020); Alon et al. (2020); Zhang et al. (2020); Chen et al. (2020); Kleinberg & Raghavan (2020); Bechavod et al. (2021); Ahmadi et al. (2022a;b); Raab & Liu (2021). Unlike strategic manipulation, improvement will incur a label change. Liu et al. (2019) studied the conditions where fairness interventions can promote improvement among individuals. Rosenfeld et al. (2020) added regularization in strategic classification algorithms to let a decision maker favor improvement. Zhang et al. (2020) formulated the label change as a transition matrix where the transition probabilities are deterministic and difficult to estimate.

Besides, several works consider both behaviors at the same time.Shavit et al. (2020) and Alon et al. (2020) introduced causal inference frameworks into strategic behaviors including manipulation and improvement. Kleinberg & Raghavan (2020) proposed a mechanism to incentivize individuals to invest on specific features where the individuals have a budget to invest strategically on all features including undesired ones. Chen et al. (2020) divided the features into immutable features, improvable features and manipulable features and explored linear classifiers which can prevent manipulation and encourage improvement. Jin et al. (2022) also focused on incentivizing improvement and proposed a subsidy mechanism to induce improvement actions and improve social well-being metrics. Bechavod et al. (2021) demonstrated the ability of strategic decision makers to distinguish features influencing the label of individuals under an online setting. Ahmadi et al. (2022a) proposed a linear model where strategic manipulation and improvement are both present. Barsotti et al. (2022) conducted several empirical experiments when both improvement and manipulation are possible where both actions incur a linear deterministic cost.

C.3 Recommendation Systems

Our work is also related to preference shifts and opinion dynamics in recommendation systems, which we refer to Castellano et al. (2009) as a comprehensive survey. Among the rich set of works, Dean & Morgenstern (2022); Gaitonde et al. (2021) proposed geometric models for opinion polarization and motivate our work.

D Illustration of Table 1



(a) x_0 to improve under (θ, r) (b) Best profile to reach for x_0

Figure 7: Illustration of Table 1

Table 1 illustrate the minimum requirement of x_0 for an individual to improve under different (θ, r), and the best attainable profile for individuals with initial similarity x_0. We illustrate them in Fig. 7.

Discussions of intervention strategies in real applications. Table 1 further suggest effective strategies that encourage individuals to improve their qualifications, i.e., more individuals are incentivized to improve if (i) the decision-maker's acceptance threshold θ is lower; or (ii) the time it takes for individuals to succeed after investments is shorter. Examples of both strategies in real applications are as follows.

1. *Lower acceptance threshold θ in hiring:* Instead of directly recruiting the qualified candidates, companies first lower the standard by offering internship opportunities to encourage applicants to improve, and then offer full-time positions. This two-stage hiring process widens the candidate pool and incentivizes more people to improve.
2. *Lower discounting factor r in college admission:* Instead of directly rejecting the unqualified high school graduates, universities incentivize them by issuing conditional transfer offers. Once these students meet certain requirements, they get admitted. The conditional acceptances encourage more students to improve by lowering the time it takes for them to receive reward.

Meanwhile, Table 1 also reveals that setting short-term goals will be effective to incentivize individuals to improve. For instance, teachers may set up several quizzes to break down the grade and make students more motivated to improve.

E Illustration of Thm. 5.1

Table 2: Ranges $(\widehat{x_1}, \widehat{x_2})$ of initial similarity x_0 under which individuals prefer to manipulate.

2*θ	2*r	Detection probability P					
		0	0.1	0.2	0.3	0.4	0.5
0.995	0.1	$(0.364, 0.995)$	$(0.435, 0.994)$	$(0.513, 0.993)$	$(0.596, 0.991)$	$(0.686, 0.984)$	$(0.796, 0.966)$
0.976	0.05	$(0.499, 0.976)$	$(0.613, 0.973)$	$(0.740, 0.958)$	$\varnothing$	$\varnothing$	$\varnothing$
0.953	0.01	$(0.773, 0.953)$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$

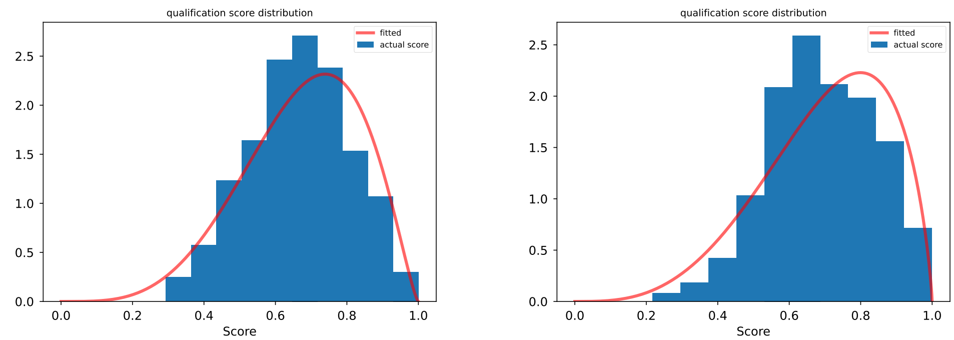


Figure 9: Exam Score: $Beta$ distributions

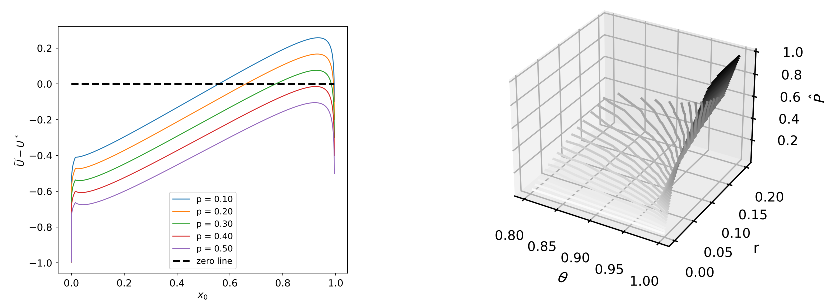


Figure 8: Illustration of Thm. 5.1: the left figure shows $\widetilde{U} - U$ as functions of x_0 under different P when $\theta = 0.995, r = 0.05$; the right plot shows threshold $\widehat{P}$ under different pairs of (θ, r).

Thm. 5.1 identifies conditions under which manipulation (or improvement) is preferred by individuals over the other. As mentioned in Section 5, the specific values of $\widehat{P}$, $\widehat{x}$, $\widehat{x_1}$, $\widehat{x_2}$ in Thm. 5.1 depend on θ, r, and we can empirically find $\widehat{P}$, $\widehat{x}$, $\widehat{x_1}$, $\widehat{x_2}$ and verify the theorem, as illustrated in Figure 8 and Table 2. Specifically, the left plot in Figure 8 shows $\widetilde{U} - U$ as functions of initial similarity x_0 under different detection probability P. Because individuals only prefer to manipulate if $\widetilde{U} - U > 0$, the plot shows the values of $\widehat{P}$, $\widehat{x}$, $\widehat{x_1}$, $\widehat{x_2}$ in Thm. 5.1. The right plot shows threshold $\widehat{P}$ under different pairs of (θ, r), and it shows that $\widehat{P}$ increases as r increases. Table 2 shows ranges $(\widehat{x_1}, \widehat{x_2})$ of initial similarity x_0 under different detection probability P, acceptance threshold θ, and discounting factor r.

F Additional Experiments

Exam Score Data

Just as Sec. 7 mentions, we acquire the exam score data Kimmons (2012), preprocess the data and fit beta distributions for both males and females. The fitted distribution and real distribution are shown in Fig. 9.

FICO Score Data

Just as Sec. 7 mentions, we fit beta distributions for FICO Score Hardt et al. (2016b), and obtain four distributions for different racial groups as shown in Fig. 10.

Additional Results for FICO Data

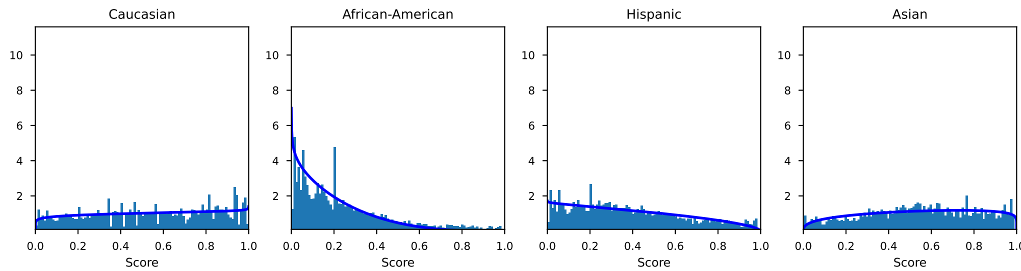


Figure 10: FICO Score: Caucasian (Beta$(1.11, 0.97)$), African American (Beta$(0.91, 3.84)$), Hispanic (Beta$(0.99, 1.58)$), Asian (Beta$(1.35, 1.13)$)

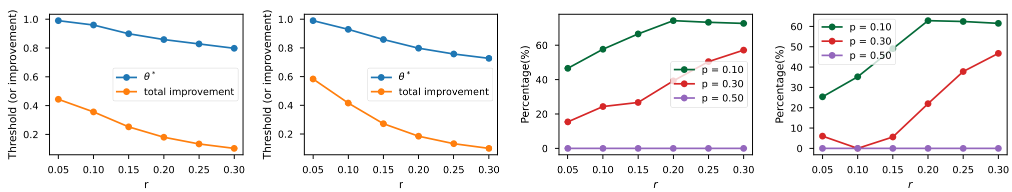


Figure 11: Optimal thresholds to incentivize improvement (left two plots) and manipulation probability under the thresholds (right two plots) for Asians and Hispanic of the **FICO data**.

Besides Caucasian and African American mentioned in Sec. 7, for Asians and Hispanic, we also compute the optimal decision threshold and corresponding total improvement under different r. As shown in Fig. 11, θ* always decreases with r and the total amount of improvement decreases. If comparing Asians and Hispanics, we observe that Hispanics have lower thresholds but larger improvements. For settings with both manipulation and improvement (Fig. 11), it seems that a larger (resp. smaller) proportion of Asians tend to manipulate than African Americans under θ^*. More importantly, the optimal thresholds reveal larger amounts of improvement for Hispanics, suggesting that the decision-maker's policy in Sec. 4 is beneficial for the disadvantaged group.

G Estimating the discounting factor r in Sec.4

We can estimate the discounting factor r if given an experimental population. The decision-maker can publish an arbitrary threshold θ and observe the lowest score among all individuals who change their scores, which is $x^*(\theta)$. Then the decision-maker can use any expression in Table 1 to estimate r. Multiple experiments can make the estimation more robust.

H Proofs

H.1 Proof Details of Thm. 3.1

To derive k^*, we first take the derivative of equation 5 with respect to k. For simplicity, let $K = k+1$ and the derivative will not change. Also, let $R = r+1$ and $G = -\ln\left(\sqrt{\frac{(\theta)^{-2}-1}{(x_0)^{-2}-1}}\right)$. Then show the results as follows:

$$\frac{\partial U}{\partial K} = \ln R \cdot R^{\frac{-G}{\ln K}} \cdot \frac{G}{K \cdot \ln^2 K} - 1 \tag{13}$$

$$\frac{\partial^2 U}{\partial K^2} = \frac{-G \cdot \ln R \cdot R^{\frac{-G}{\ln K}} (\ln^2 K + 2\ln K - G \cdot \ln R)}{K^2 \cdot \ln^4 K} \tag{14}$$

The denominator of $\frac{\partial^2 U}{\partial K^2}$ is always positive, and the first term $-G \cdot \ln R \cdot R^{\frac{-G}{\ln K}}$ of numerator is always negative.

Also, because $K \in [1, 2]$, $\ln^2 K + 2 \ln K \in (0, \ln^2 2 + 2 \ln 2)$. Thus, we have following situations:

1) If $G \cdot \ln R > \ln^2 2 + 2 \ln 2$, $\frac{\partial^2 U}{\partial K^2}$ is always positive when $K \in [1, 2]$. This means $\frac{\partial U}{\partial K}$ is increasing.

Then, noticing that $\lim_{K \to 1^+} \frac{\partial U}{\partial K} = -1$, we know $\frac{\partial U}{\partial K}$ is always negative when $K \in [1, 2]$. This means U is monotonically decreasing. Also, when $k = 0$, $U = 0$. This ensures U is always non-positive and individuals will never choose to invest any effort.

2) If $G \cdot \ln R \leqslant \ln^2 2 + 2 \ln 2$, $\frac{\partial^2 U}{\partial K^2}$ is first positive, then negative when $K \in [1, 2]$. Also, if plugging $K = 2$ into equation 13, we know $\lim_{K \to 2} \frac{\partial U}{\partial K} < 0$. These facts reveal that $\frac{\partial U}{\partial K}$ is firstly increasing from a negative number and then decreasing to a negative number. And there must exist a unique maximum point when $K = K^{'}$, $K^{'}$ should satisfy:

$$\ln^2 K^{'} + 2 \ln K^{'} - G \cdot \ln R = 0 \tag{15}$$

Plug equation 15 into equation 13. Denote $\ln K^{'}$ as $t \in [0, \ln 2]$, and denote $\frac{\partial U}{\partial K}$ at $K^{'}$ as L:

$$L = \frac{t + 2}{t \cdot e^{2t+2}} - 1 \tag{16}$$

Then take the derivative of L:

$$\frac{\partial L}{\partial t} = \frac{-2(t + 1)^2 \cdot e^{2t+2}}{t^2 \cdot e^{4t+4}} < 0 \tag{17}$$

equation 17 shows L is decreasing. Also, noticing that $\lim_{t \to 0^+} L(t) = +\infty$ and $\lim_{t \to \ln 2} L(t) < \frac{3}{2e^2} - 1 < 0$, we know there must exist a $t^{'} \in (0, \ln 2)$ as the root of M. We can explicitly solve $t^{'} = 0.1997$.

Thus, we now know that when $t \in [0, t^{'}]$, $L \geqslant 0$. With the plausible domain of t and equation 15, we would know: When $G \ln R \in [0, t^{'2} + 2t^{'}]$, $L \geqslant 0$ and thereby U has an extreme large point with value $U*$. At this maximum point, equation 13 equals 0, and equation 14 is smaller than 0.

Finally, we derive the condition for $U* > 0$: Denote $G \ln R$ as C and $\ln K$ as z, U can be simplified to:

$$U = e^{\frac{-C}{z}} - e^z + 1 \tag{18}$$

Because $z \in [0, \ln 2]$, for any t fixed, $\lim_{C \to 0} U = 2 - e^z \geqslant 0$ and $\lim_{C \to 0} U = 1 - e^z \leqslant 0$. With the fact that $\frac{\partial U}{\partial C} < 0$, we know U is monotonically decreasing with C, so is $U*$. Thus, there must exist a threshold m, when $C < m$, $U* > 0$. And if $U* > 0$, individuals will decide to improve. Then Thm. 3.1 is proved and we can numerically solve the threshold $m = 0.316$.

Although we believe exponential discounting is general and fits our setting well, we also note that we can still use derivative analysis when the discounting changes (e.g., hyperbolic discounting). Specifically, if denoting the discounted reward as $d(r, t)$, we would have $U = d(r, T) - k$. Then if taking the derivative we will get $\frac{\partial U}{\partial k} = \frac{\partial d}{\partial T} \cdot \frac{\partial T}{\partial k} - 1$. Noticing that T is known, then discussing the properties of d with different choices of discounting is enough to derive the nature of U.

H.2 PROOF DETAILS OF THM. 4.1 AND COROLLARY 4.2

H.2.1 PROOF OF THM. 4.1

First prove $U_d(\theta)$ has a maximize $\theta* \in (0, 1)$:

With the definition of $U_d(\theta)$ in equation 7, we already know U_d is continuous. We can first observe that $U_d(0) = 0, U_d(1) = 1$. These hold simply because $x^*(0) = 0 and x^*(1) = 1$. Next noticing that for any $\theta \in (0, 1)$, $U_d(\theta) > 0$ holds. This suggests that θ will reach its maximum point according to the Weierstrass extreme value theorem.

Next, noticing that $U_d(\theta) > 0 \in (0, 1)$ we can derive that $\frac{\partial U_d}{\partial \theta}(0) > 0$ and $\frac{\partial U_d}{\partial \theta}(0) < 0$. Then if it only has one root in $(0, 1)$, we would know U_d must first increase and then decrease because there is at most one inflection point. Thus, a unique maximum exists.

H.2.2 Proofs of why Uniform distribution has a unique maximized θ^*

If $\frac{\partial U_d}{\partial \theta}$ only has one root. We know it is first larger than 0, then becomes smaller than 0. Next, according to the Leibniz integral rule, we can get:

$$\frac{\partial U_d}{\partial \theta} = \int_{x^*(\theta)}^{\theta} P(x)dx - (\theta - x^*(\theta)) \cdot P(x^*(\theta)) \cdot \frac{\partial x^*(\theta)}{\partial \theta}$$

Use Lagrange's Mean Value Theorem, we can write the above equation as:

$$(\theta - x^*(\theta)) \cdot [P(\theta^{'}) - P(x^*(\theta)) \cdot \frac{\partial x^*(\theta)}{\partial \theta}]$$

where $\theta^{'}$ is between $x^*(\theta), \theta$. Thus, the second term $P(\theta^{'}) - P(x^*(\theta)) \cdot \frac{\partial x^*(\theta)}{\partial \theta}$ must also be first larger than 0 then smaller than 0. Next, noticing that $P(\theta^{'}) = P(x^*(\theta))$ in uniform distribution and $\frac{\partial x^*(\theta)}{\partial \theta}$ is increasing, the equation will be smaller than 0 when $\frac{\partial x^*(\theta)}{\partial \theta} < 1$ and vice versa. Thus, we prove the result for the uniform distribution.

H.2.3 Proof of Corollary 4.2

We now know $\frac{\partial U_d}{\partial \theta} = (\theta - x^*(\theta)) \cdot [P(\theta^{'}) - P(x^*(\theta))] \cdot \frac{\partial x^*(\theta)}{\partial \theta}]$. Then according to the expression of $x^*(\theta)$, it is true that both $x^*(\theta)$ and $\frac{\partial x^*(\theta)}{\partial \theta}$ increase with r. Thus, when the probability distribution remains unchanged, the root of $\frac{\partial U_d}{\partial \theta}$ when r increases becomes smaller.

H.3 Proof Details of Thm. 5.1

Denote $\ln\left(\sqrt{\frac{\theta^{-2}-1}{x_0^{-2}-1}}\right)$ as $G(x_0)$. $G(x_0)$ is always negative and monotonically increasing with $x_0 \in (0, \theta)$.

1. Situation when P = 0

According to Sec. 3 and equation 8, we can write the maximum improvement utility U^* as $(1 + r)^{\frac{G(x_0)}{ln(k^*+1)}} - k^*$, and write manipulation utility $\widetilde{U}$ as $(1 + r)^{\frac{G(x_0)}{ln2}} - (\theta - x_0)$.

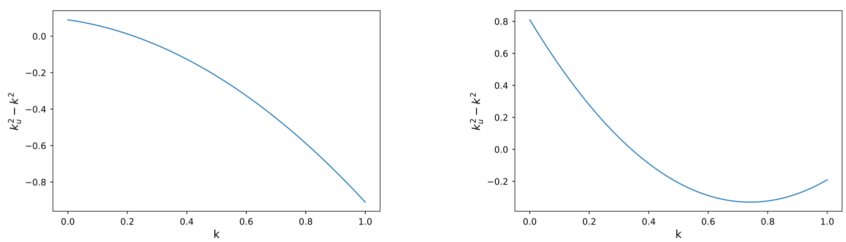
Then take the derivative of both:

$$\frac{\partial U^*}{\partial x_0} \geqslant \frac{\partial G}{\partial x_0} \cdot \frac{ln(1 + r)}{ln(k^* + 1)} \cdot (1 + r)^{\frac{G(x_0)}{ln(k^*+1)}} \tag{19}$$

$$\frac{\partial \widetilde{U}}{\partial x_0} = \frac{\partial G}{\partial x_0} \cdot \frac{ln(1 + r)}{ln2} \cdot (1 + r)^{\frac{G(x_0)}{ln2}} + 1 \tag{20}$$

The "$\geqslant$" in equation 19 occurs because k^* is actually a function of x_0, but if we regard k^* at x_0 as a constant, the derivative here serves as a lower bound of $\frac{\partial U^*}{\partial x_0}$.

Firstly, we prove when $x_0 \to \theta, U^* < \widetilde{U}$: when $x_0 \to \theta$, we know $k^* \to 0$ since individuals invest an arbitrarily small effort to immediately qualified. However, according to Sec. H.1, k^* should let



(a) shape when $2x_0^2 + 2x_0^3 - 1 < 0$ (b) shape when $2x_0^2 + 2x_0^3 - 1 > 0$

Figure 13: Shapes of $k_u^2 - k^2$

$\frac{\partial^2 U}{\partial k^2} < 0$. This inequality will give us the bound of k^*: $\ln(k^* + 1) > \frac{-G(x_0) \cdot ln(1+r)}{3}$. With this bound, we can plug k^* into equation 19, and know $\frac{ln(1+r)}{ln(k^*+1)} \to +\infty$, and $(1+r)^{\frac{G(x_0)}{ln(k^*+1)}}$ is larger than a constant because of the bound. Therefore, $\frac{\partial U^*}{\partial x_0} \geqslant \frac{\partial G}{\partial x_0} \cdot +\infty$. Then according to equation 20, when $x_0 \to \theta$, $\frac{\partial \widetilde{U}}{\partial x_0} < \frac{\partial G}{\partial x_0} \cdot \frac{ln(1+r)}{ln2} + 1$. Since $\frac{\partial G}{\partial x_0}$ is always positive, when $x_0 \to \theta$, we prove that $\frac{\partial U^*}{\partial x_0} > \frac{\partial \widetilde{U}}{\partial x_0}$. Meanwhile, when $x_0 = \theta$, $U^* = \widetilde{U} = 1$. This means when $x_0 \to \theta$, $U^* < \widetilde{U}$.

Secondly, when $x_0 = 0$: $\widetilde{U} = -\theta$ and $U^* = 0$. So $\widetilde{U} < U^*$ when $x_0 = 0$.

Thus. there must be an intersection between $\widetilde{U}$ and U^*. Then noticing that if we increase θ, $\widetilde{U}$ is always decreasing to converge to function $y = x - 1$, while $U^* \geqslant 0$ always holds. This suggests when θ is sufficiently close to 1, we can guarantee the first intersection of U^* and $\widetilde{U}$ occurs arbitrarily close to 1, meaning this first intersection is the only intersection.

Let the only *intersection* be $\hat{x}$, we prove situation 1. The shapes of $\widetilde{U}$ and U^* are illustrated in Fig. 12.

2. Situation when P > 0

From equation 8: when $x_0 \to \theta$, $\widetilde{U} \to 1 - P$. However, at this time $U^* \to 1 > 1 - P$. This demonstrates $\hat{x_2}$ must exist.

When $P \to 0$, according to situation 1 and the continuity of $\widetilde{U}$ with respect to P, $\hat{x_1}$ must exist. However, when $P \to 1$, $\widetilde{U}$ is always negative, making $\hat{x_1}$ does not exist.

Thus, there must exist a threshold $\hat{P}$, when $P \leqslant \hat{P}$, $\hat{x_1}, \hat{x_2}$ exist. Otherwise, $U^* > \widetilde{U}$ is always true.

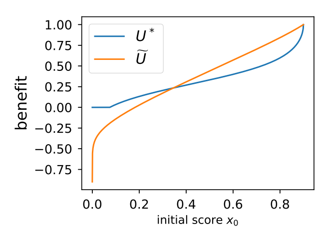


Figure 12: $\widetilde{U}, U^*$

H.4 Proof Details of Thm. 6.2

First let us prove following two lemmas:

Lemma H.1. *For any initial qualification score x_0, There exists a $\widehat{k} \in (0,1)$, when $k \in [0, \widehat{k})$, $k_u > k$. Let $\widehat{x_0}$ be the only root of $2x_0^2 + 2x_0^3 - 1 = 0$ within $(0,1)$, then $\widehat{k}$ is given by:*

$$\widehat{k} = \min(\frac{\widehat{x_0}^2}{2\widehat{x_0}^2 + 2\widehat{x_0}^3}, \frac{x_0 \cdot (x_0^2 + x_0 - \sqrt{x_0^4 - x_0^2 + 1})}{2x_0^2 + 2x_0^3 - 1}) \tag{21}$$

Proof

According to Thm. 6.1, $k_u^2 = \|\widetilde{d}\|^2 \cdot x_0^2$ and $\|\widetilde{d}\|^2 = k^2 + (1-k)^2 - 2k(1-k)x_0$. We can get following expression:

$$k_u^2 - k^2 = (2x_0^2 + 2x_0^3 - 1)k^2 - (2x_0^2 + 2x_0^3)k + x_0^2 \tag{22}$$

Firstly, when $2x_0^2 + 2x_0^3 - 1 = 0$, $\widehat{x_0} = 0.565$. Thus, when $k < \frac{\widehat{x_0}^2}{2\widehat{x_0}^2 + 2\widehat{x_0}^3} = 0.319$, $k_u^2 > k^2$.

Except the above situation, We can regard equation 22 as a quadratic function of k and solve the two roots:

$$\frac{x_0 \cdot (x_0^2 + x_0 \pm \sqrt{x_0^4 - x_0^2 + 1})}{2x_0^2 + 2x_0^3 - 1} \tag{23}$$

We then prove a claim that when $x_0 \in (0, 1)$, $\frac{x_0 \cdot (x_0^2 + x_0 + \sqrt{x_0^4 - x_0^2 + 1})}{2x_0^2 + 2x_0^3 - 1}$ is either larger than 1 or smaller than 0:

1) When $2x_0^2 + 2x_0^3 - 1 < 0$, the denominator of equation 23 is negative, while the numerator is always positive. Thus, equation 23 is negative.

2) When $2x_0^2 + 2x_0^3 - 1 > 0$:

$$\frac{x_0 \cdot (x_0^2 + x_0 + \sqrt{x_0^4 - x_0^2 + 1})}{2x_0^2 + 2x_0^3 - 1} > \frac{x_0 \cdot (x_0^2 + x_0 + x_0^2)}{2x_0^3 + x_0^2} = 1 \tag{24}$$

equation 24 means equation 23 is larger than 1. Thus, the claim is proved.

Thus, $k_u^2 - k^2$ only has one root within $(0, 1)$. Also from equation 22 we know when $k = 0$, $k_u > k$ and when $k = 1$, $k_u \leqslant k$. With these facts we immediately know: When $k \leqslant \frac{x_0 \cdot (x_0^2 + x_0 - \sqrt{x_0^4 - x_0^2 + 1})}{2x_0^2 + 2x_0^3 - 1}$, $k_u^2 - k^2 \geqslant 0$. Otherwise, $k_u^2 - k^2 < 0$. In fact, besides the exception $2x_0^2 + 2x_0^3 - 1 = 0$, there are only two possibilities of the shape of $k_u^2 - k^2$ as shown in Fig. 13. Because k and k_u are both non-negative, the relationship of the square must be the same for their values.

Then if we define $\hat{k}$ as:

$$\hat{k} = \min\left(\frac{\widehat{x_0}^2}{2\widehat{x_0}^2 + 2\widehat{x_0}^3}, \frac{x_0 \cdot (x_0^2 + x_0 - \sqrt{x_0^4 - x_0^2 + 1})}{2x_0^2 + 2x_0^3 - 1}\right) \tag{25}$$

Then $k_u > k$ when $k \in [0, \hat{k})$. Proved.

Lemma H.2. *For any individual with initial qualification score x_0 and the admission threshold θ, there must exist a r to let there exists a $\bar{k} \in [0, \widehat{k})$, $U(\bar{k}, \theta, r, x_0) > 0$*

Proof. If we let $z = \ln(k + 1)$ be z and recall that $C(\theta, x_0, r) = -\ln\left(\sqrt{\frac{(\theta)^{-2} - 1}{(x_0)^{-2} - 1}}\right) \cdot \ln(1 + r)$, we would have $U = e^{\frac{-C}{z}} - e^z + 1$.

For any z there exists C_z, when $C < C_z$, $U > 0$.

So we can just let k be an arbitrary point $\in [0, \widehat{k})$ and we can get the corresponding C_z, then we can only let r satisfy:

$$\ln(1 + r) < \frac{C_z}{-\ln\left(\sqrt{\frac{(\theta)^{-2} - 1}{(x_0)^{-2} - 1}}\right)} \tag{26}$$

Then we find the plausible r. Proved.

Proof of Thm. 6.2

According to Lemma H.1, when $\bar{k} \in [0, \widehat{k})$, $k_u > k$, so the convergence speed of the individual to d^* under forgetting mechanism will be faster than the convergence speed of the individual to d without forgetting mechanism, so that the reward under forgetting mechanism is discounting less. Meanwhile, according to Lemma H.2, there exists a r where $U(\bar{k}, \theta, r, x_0) > 0$. Combine them together, $\widehat{U}(\bar{k}, \theta, r, x_0) > U(\bar{k}, \theta, r, x_0) > 0$ and Thm. 6.2 is proved.

H.5 PROOF OF THM. B.1

Assume $k_t = \frac{k}{t+1}$ when $t \geqslant 0$. From equation 1 and similar to Dean & Morgenstern (2022), we know $(q_{t+1}^T \cdot d)^{-2} - 1 = \frac{(q_t^T \cdot d)^{-2} - 1}{k_t + 1}^2$. This will lead to $(q_t^T \cdot d)^{-2} - 1 = \prod_{i=0}^{t-1} (\frac{k}{i+1} + 1)^{-2} ((q_0^T \cdot d)^{-2} - 1)$.

Then consider $\prod_{i=0}^{t-1} (\frac{k}{i+1} + 1)^{-1} = \prod_{i=0}^{t-1} (\frac{i+1}{k+i+1}) = \frac{1}{k+1} \cdot \frac{2}{k+2} \dots$. When $k = 1$, The expression is $\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \dots \frac{t-1}{t} = \frac{1}{t}$, demonstrating the convergence rate is linear. Note that this expression is decreasing as k decreases, so the convergence rate in our model is always slower than linear. Next, consider the general expression $\prod_{i=0}^{t-1} (\frac{i+1}{k+i+1}) = \frac{1}{k+1} \cdot \frac{2}{k+2} \dots$ and $k < 1$. Let $a = \frac{1}{k}$ which is larger than 1, and $j = i + 1$ which is larger than 0. We slightly abuse the definition of a to let it be an integer. Then the expression becomes $\prod_{i=0}^{t-1} (\frac{ja}{1+ja}) = \frac{a}{a+1} \cdot \frac{2a}{2a+1} \dots \frac{ta}{ta+1}$.

Then for any a we can bound this expression. Basically, we already know $\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \dots \frac{t-1}{t} = \frac{1}{t}$. Noticing that when $a > 1$, it is just equal to erase some terms of this expression. We can utilize this fact to get the lower bound and upper bound:

1. Lower bound: consider the following $a-1$ sets of expressions and each set consists of t terms: $\{\frac{1}{2} \cdot \frac{a+1}{a+2} \cdot \frac{2a+1}{2a+2} \dots \cdot \frac{(t-1)a+1}{(t-1)a+2}\}, \{\frac{2}{3} \cdot \frac{a+2}{a+3} \cdot \frac{2a+2}{2a+3} \dots \cdot \frac{(t-1)a+2}{(t-1)a+3}\} \dots, \{\frac{a-1}{a} \cdot \frac{2a-1}{2a} \cdot \frac{3a-1}{3a} \dots \cdot \frac{ta-1}{ta}\}$. Then each of the $a-1$ expressions are smaller than $\prod_{j=1}^{t} (\frac{ja}{1+ja}) = \frac{a}{a+1} \cdot \frac{2a}{2a+1} \dots \frac{ta}{ta+1}$. Denote $\prod_{j=1}^{t} (\frac{ja}{1+ja}) = \frac{a}{a+1} \cdot \frac{2a}{2a+1} \dots \frac{ta}{ta+1}$ as I, we will have $I^a \geqslant \frac{1}{ta+1}$, so the convergence rate is smaller than $\sqrt[a]{ta} = \Theta(t^k)$
2. Upper bound: consider the following $a-1$ sets of expressions and each set consists of t terms: $\{\frac{a+1}{a+2} \cdot \frac{2a+1}{2a+2} \dots \cdot \frac{ta+1}{ta+2}\}, \{\frac{a+2}{a+3} \cdot \frac{2a+2}{2a+3} \dots \cdot \frac{ta+2}{ta+3}\} \dots, \{\frac{2a-1}{2a} \cdot \frac{3a-1}{3a} \dots \cdot \frac{(t+1)a-1}{(t+1)a}\}$. Then each of the $a-1$ expressions are larger than $\prod_{j=1}^{t} (\frac{ja}{1+ja}) = \frac{a}{a+1} \cdot \frac{2a}{2a+1} \dots \frac{ta}{ta+1}$. Denote $\prod_{j=1}^{t} (\frac{ja}{1+ja}) = \frac{a}{a+1} \cdot \frac{2a}{2a+1} \dots \frac{ta}{ta+1}$ as I, we will have $I^a \leqslant \frac{1}{(t+1)}$, so the convergence rate is larger than $\sqrt[a]{ta} = \Theta(t^k)$

Thus, take the limit and apply the Sandwich Theorem, the convergence rate is $\Theta(t^k)$.