When Machines Pass the Test: Professional Certification Signaling Erosion Under AI Disruption

Wei-Lun Cheng[a], Daniel Wei-Chung Miao[a,*], Guang-Di Chang[a]

[a]*Graduate Institute of Finance, National Taiwan University of Science and Technology, Taipei, 10607, Taiwan*

Abstract

Professional certifications such as the Chartered Financial Analyst (CFA) designation have long served as credible signals in financial labor markets, enabling employers to screen for scarce cognitive abilities. We develop a Modified Spence Signaling Model that incorporates an AI replication cost parameter to analyze how large language models (LLMs) erode certification signaling value. Our central theoretical result—the Partial Signaling Collapse Theorem—demonstrates that signaling erosion is *selective*, not total: certification loses its screening power over formalizable abilities (formula recall, rule application, algorithmic computation) where AI replication costs approach zero, but retains signaling value for tacit abilities (ethical judgment, stakeholder reasoning, fiduciary decision-making) that resist low-cost AI replication. Integrating the Autor et al. [3] task-based framework with Becker [4] human capital theory, we map the CFA curriculum onto a six-dimensional ability taxonomy and derive a tipping-point condition: when the fraction of AI-replicable abilities exceeds a critical threshold α^*, the separating equilibrium collapses into pooling. We establish that the CFA program, in its current form, is approaching this threshold. Our analysis yields concrete policy implications: certification bodies must rebalance assessment toward AI-resistant competencies—context-dependent judgment, interpersonal reasoning, and responsibility-bearing decision-making—to preserve institutional credibility in the age of artificial intelligence.

*Corresponding author

Email addresses: `d11018003@mail.ntust.edu.tw` (Wei-Lun Cheng), `miao@mail.ntust.edu.tw` (Daniel Wei-Chung Miao), `gchang@mail.ntust.edu.tw` (Guang-Di Chang)

---

1. Introduction

The Chartered Financial Analyst (CFA) designation has served as one of the most widely recognized professional certifications in global finance for over six decades. With a historically low pass rate—averaging approximately 43% for Level I and declining to roughly 50% for Level III—the certification imposes substantial costs on candidates in terms of time, effort, and foregone earnings [7]. In the classical labor economics framework of Spence [11], these costs are precisely what make the CFA credential a credible signal: because high-ability workers find it less costly to obtain the certification, the CFA charter sustains a separating equilibrium in which employers can distinguish between ability types.

The rapid advancement of large language models (LLMs) poses a fundamental challenge to this signaling mechanism. Recent studies demonstrate that frontier AI systems achieve performance levels on standardized financial examinations that rival or exceed median human candidates. Callanan et al. [6] show that GPT-4 passes the CFA Level I and Level II examinations, while domain-adapted models such as BloombergGPT [13] and FinDAP's Llama-Fin [9] demonstrate strong performance on financial knowledge benchmarks. When an AI system can replicate the cognitive skills that a certification is designed to measure—at near-zero marginal cost—the fundamental economic logic of signaling is disrupted.

This paper asks a precise question: *Does AI replication of certified cognitive abilities destroy the signaling value of professional certification, and if so, how?* We argue that the answer is neither a simple yes nor a simple no. Instead, we develop a theoretical framework demonstrating that signaling erosion is *partial and selective*—concentrated on formalizable abilities while sparing tacit competencies that resist cheap AI replication.

Our contribution is threefold. First, we extend the Spence [11] signaling model by introducing a multi-dimensional ability space and an AI replication cost function, deriving equilibrium conditions under which certification retains or loses signaling power. Second, we integrate the Autor et al. [3] task-based framework with Becker [4] human capital theory to provide a taxonomy of professional abilities ranked by AI vulnerability. Third, we apply

this framework to the CFA certification specifically, mapping its three-level curriculum onto our ability taxonomy and deriving concrete policy implications for assessment design.

The remainder of this paper is organized as follows. Section 2 reviews the relevant theoretical literature. Section 3 develops the Modified Spence Signaling Model. Section 4 derives the partial signaling collapse result and tipping-point conditions. Section 5 applies the framework to the CFA certification. Section 6 discusses empirical implications and policy recommendations. Section 7 addresses extensions and limitations. Section 8 concludes.

2. Theoretical Foundations

2.1. Signaling in Labor Markets

The canonical signaling model of Spence [11] establishes that in markets with asymmetric information, workers invest in observable signals—such as education or professional certification—to convey their unobservable ability to employers. The model's central insight is that signaling is credible only when the cost of acquiring the signal is negatively correlated with ability: high-ability workers find it cheaper to obtain the credential, sustaining a *separating equilibrium* in which the signal is informative.

Subsequent work has extended this framework in several directions. Riley [10] surveys twenty-five years of signaling theory development. Tyler et al. [12] provides empirical evidence on the signaling value of educational credentials, while Bedard [5] examines how changes in the cost structure of education affect equilibrium outcomes. A consistent finding is that signals lose value when the cost differential between types shrinks—precisely the dynamic that AI replication introduces.

2.2. Human Capital Theory and Professional Certification

Becker [4] distinguishes between general human capital (transferable across firms) and specific human capital (valuable only within particular contexts). Professional certifications like the CFA primarily certify general human capital: knowledge and skills that are portable across employers in the financial industry. This generality is precisely what makes certifications valuable as market-wide signals—but also what makes them vulnerable to AI disruption, since general knowledge is by definition *codifiable* and thus more amenable to machine replication.

3

The human capital framework suggests a compositional shift rather than wholesale destruction: AI does not eliminate human capital but alters the relative market value of its components. General cognitive skills (formula recall, rule application) depreciate as AI provides them cheaply, while specific and tacit skills (contextual judgment, relationship management) appreciate in relative terms [8].

2.3. The Task-Based Framework

Autor et al. [3] propose a taxonomy of workplace tasks that has become foundational for analyzing technological displacement:

- **Routine cognitive tasks**: rule-based mental activities that follow explicit procedures (e.g., bookkeeping, formulaic calculation).
- **Routine manual tasks**: repetitive physical activities.
- **Non-routine analytical tasks**: problem-solving requiring pattern recognition and structured reasoning.
- **Non-routine interactive tasks**: activities requiring negotiation, persuasion, mentoring, and contextual judgment.

Their central result is that computerization substitutes for routine tasks while complementing non-routine tasks. Acemoglu and Restrepo [1] extend this to a general equilibrium framework, modeling automation as the assignment of tasks previously performed by humans to machines, with implications for wages, employment, and the creation of new comparative advantages for human labor.

We observe that large language models represent a qualitative shift in this framework: unlike earlier automation technologies, LLMs can perform certain *non-routine analytical tasks* that were previously considered resistant to computerization. This expansion of the "automatable frontier" is precisely what generates novel implications for professional certification signaling.

2.4. AI and Financial Expertise

A growing empirical literature documents AI performance on financial professional tasks. Callanan et al. [6] evaluate GPT-4 on CFA examinations, finding pass-level performance on Levels I and II. Wu et al. [13] develop a domain-specific 50-billion parameter model trained on financial corpora, demonstrating strong performance on financial NLP benchmarks. Ke et al. [9] introduce the FinDAP framework for domain-adaptive post-training, adapting Llama-3-8B to the financial domain through a three-stage pipeline

(continual pre-training, supervised fine-tuning, and preference alignment), achieving substantial gains on financial reasoning benchmarks.

These results motivate our theoretical investigation: if the cognitive abilities certified by professional examinations can be replicated by AI systems at low cost, what are the equilibrium implications for the labor market role of certification?

3. A Modified Spence Signaling Model with AI Replication

3.1. Setup

Consider a labor market with asymmetric information. There is a continuum of workers, each characterized by an unobservable ability type $\theta \in \{\theta_L, \theta_H\}$, where $\theta_H > \theta_L > 0$. The prior probability that a worker is high-ability is $\lambda \in (0, 1)$. Workers can invest in a professional certification $e \in \{0, 1\}$ (e.g., the CFA charter) at cost $c(\theta)$, where the single-crossing property holds: $c(\theta_H) < c(\theta_L)$.

We depart from the standard model by introducing a *multi-dimensional ability space.* Let certification e attest to a vector of K distinct skill dimensions:

$$\mathbf{s} = (s_1, s_2, \ldots, s_K) \in \mathbb{R}_+^K \tag{1}$$

where each s_k represents a specific cognitive ability that the certification is designed to measure and signal. In the context of the CFA, these dimensions include declarative knowledge (s_1), algorithmic computation (s_2), analytical decomposition (s_3), integrative judgment (s_4), normative/ethical reasoning (s_5), and stakeholder reasoning (s_6).

Assumption 1 (Employer Valuation). *The employer's valuation of a certified worker is a weighted sum over skill dimensions:*

$$V(\mathbf{s}) = \sum_{k=1}^{K} w_k \cdot s_k, \quad where \sum_{k=1}^{K} w_k = 1, \ w_k > 0 \ \forall k. \tag{2}$$

The weights w_k reflect the market's assessment of each skill's contribution to worker productivity. In the pre-AI equilibrium, these weights are relatively stable and determined by the technology of production.

3.2. The AI Replication Cost Function

The key innovation of our model is the introduction of an *AI replication cost function* that captures the cost at which an artificial intelligence system can replicate each skill dimension.

Definition 1 (AI Replication Cost). *For each skill dimension s_k, define the AI replication cost $c_{AI}(s_k) \geq 0$ as the marginal cost at which an AI system can produce output of equivalent quality to a human worker possessing ability s_k. The vector of AI replication costs is:*

$$\mathbf{c}_{AI} = \big(c_{AI}(s_1), c_{AI}(s_2), \dots, c_{AI}(s_K)\big). \tag{3}$$

Assumption 2 (Heterogeneous Replicability). *AI replication costs are heterogeneous across skill dimensions. Specifically, for the CFA ability space, there exists a partition $\{1, \dots, K\} = \mathcal{F} \cup \mathcal{T}$ into* formalizable *abilities ($\mathcal{F}$) and* tacit *abilities ($\mathcal{T}$) such that:*

$$c_{AI}(s_k) \to 0 \ \ \forall k \in \mathcal{F}, \qquad c_{AI}(s_k) \gg 0 \ \ \forall k \in \mathcal{T}. \tag{4}$$

This partition follows directly from the Autor et al. [3] task-based framework. Formalizable abilities ($\mathcal{F}$) correspond to routine cognitive tasks and the subset of non-routine analytical tasks that are amenable to pattern-based replication by LLMs. Tacit abilities ($\mathcal{T}$) correspond to non-routine interactive tasks that require embodied judgment, ethical deliberation, and contextual reasoning that resists low-cost AI replication.

3.3. AI Replicability Index

To formalize the degree of AI vulnerability for each skill dimension, we define:

Definition 2 (AI Replicability). *The AI replicability of skill dimension s_k is:*

$$\rho_k = 1 - \frac{c_{AI}(s_k)}{\bar{c}_k} \tag{5}$$

where $\bar{c}_k$ is the human acquisition cost of skill s_k (i.e., the investment required for a human to develop this ability through training and education). When $\rho_k \to 1$, the skill is fully replicable by AI; when $\rho_k \to 0$, the skill remains a human comparative advantage.

3.4. Signaling Value Under AI Replication

In the classical Spence model, certification generates signaling value because it credibly communicates unobservable ability. AI replication disrupts this mechanism by providing an alternative, low-cost source of certified-equivalent output. We formalize this as follows.

Definition 3 (Effective Signaling Value). *The effective signaling value of certification e under AI replication is:*

$$\Sigma(\mathbf{s}, \mathbf{c}_{AI}) = \sum_{k=1}^{K} w_k \cdot s_k \cdot \left(1 - \rho_k\right) = \sum_{k=1}^{K} w_k \cdot s_k \cdot \frac{c_{AI}(s_k)}{\bar{c}_k} \tag{6}$$

The intuition is direct: the signaling contribution of skill s_k is discounted by the factor $(1 - \rho_k)$, which captures the residual scarcity of that skill after accounting for AI replication. When $\rho_k = 1$, the skill contributes zero signaling value because AI provides it at negligible cost; when $\rho_k = 0$, the full signaling value is preserved.

3.5. Modified Employer Beliefs

In the pre-AI regime, an employer observing certification $e = 1$ updates beliefs according to:

$$\mathbb{E}[\theta \mid e = 1] = \theta_H \cdot \Pr(\theta_H \mid e = 1) + \theta_L \cdot \Pr(\theta_L \mid e = 1) \tag{7}$$

Under AI disruption, the employer must additionally consider whether the certified skills *continue to differentiate* the worker from an AI-augmented uncertified worker. We define the employer's *residual information gain* from observing $e = 1$ as:

$$\Delta I(e) = \sum_{k=1}^{K} w_k \cdot \left[\mathbb{E}[s_k \mid e = 1, \theta_H] - \mathbb{E}[s_k \mid \text{AI}]\right] \cdot (1 - \rho_k) \tag{8}$$

When $\rho_k \to 1$ for skill k, the term $(1 - \rho_k) \to 0$, and that skill dimension no longer contributes to the employer's information gain from observing the certification signal. The certification becomes informationally *equivalent to noise* along that dimension.

4. Equilibrium Analysis: Partial Signaling Collapse

4.1. Pre-AI Separating Equilibrium

In the standard Spence [11] framework, a separating equilibrium exists when the following incentive compatibility constraints are satisfied:

$$\text{High type:} \quad V(\mathbf{s}) - c(\theta_H) \geq V_0 \tag{9}$$

$$\text{Low type:} \quad V_0 \geq V(\mathbf{s}) - c(\theta_L) \tag{10}$$

where V_0 is the market wage for uncertified workers. The separating equilibrium requires $c(\theta_H) < V(\mathbf{s}) - V_0 < c(\theta_L)$: the wage premium from certification exceeds the cost for high types but not for low types.

4.2. AI-Modified Equilibrium Conditions

Under AI replication, the employer's willingness to pay for certified skills changes. The effective wage premium from certification becomes:

$$\Pi_{\text{AI}} = \Sigma(\mathbf{s}, \mathbf{c}_{\text{AI}}) - V_0 = \sum_{k=1}^{K} w_k \cdot s_k \cdot (1 - \rho_k) - V_0 \tag{11}$$

The modified incentive compatibility constraints are:

$$\text{High type:} \quad \Pi_{\text{AI}} \geq c(\theta_H) \tag{12}$$

$$\text{Low type:} \quad c(\theta_L) \geq \Pi_{\text{AI}} \tag{13}$$

Remark 1. *As $\rho_k \to 1$ for an increasing number of skill dimensions, Π_{AI} declines monotonically. The separating equilibrium is sustained only as long as $\Pi_{AI} > c(\theta_H)$. When AI replication costs fall sufficiently across enough dimensions, the wage premium drops below the cost of certification even for high types, and the separating equilibrium collapses.*

4.3. The Partial Signaling Collapse Theorem

We now state the central theoretical result of this paper.

Proposition 1 (Partial Signaling Collapse). *Let $\mathcal{F}$ and $\mathcal{T}$ denote the sets of formalizable and tacit skill dimensions, respectively, with $|\mathcal{F}| + |\mathcal{T}| = K$. Suppose AI replication costs satisfy Assumption 2: $\rho_k \to 1$ for all $k \in \mathcal{F}$ and $\rho_k \approx 0$ for all $k \in \mathcal{T}$. Then the effective signaling value converges to:*

$$\Sigma(\mathbf{s}, \mathbf{c}_{AI}) \ \to \ \Sigma_{\mathcal{T}} \equiv \sum_{k \in \mathcal{T}} w_k \cdot s_k \tag{14}$$

That is, the signaling value of certification collapses to the weighted sum of tacit abilities only. Signaling erosion is partial*: it eliminates the informational content of formalizable skills while preserving the signaling value of tacit skills.*

Proof. By Definition 3, the effective signaling value is:

$$\Sigma(\mathbf{s}, \mathbf{c}_{\mathrm{AI}}) = \sum_{k \in \mathcal{F}} w_k s_k (1 - \rho_k) + \sum_{k \in \mathcal{T}} w_k s_k (1 - \rho_k).$$

Under Assumption 2, $\rho_k \to 1$ for $k \in \mathcal{F}$, so $(1 - \rho_k) \to 0$ for these dimensions. Simultaneously, $\rho_k \approx 0$ for $k \in \mathcal{T}$, so $(1 - \rho_k) \approx 1$. Taking the limit:

$$\lim_{\rho_k \to 1, \, k \in \mathcal{F}} \Sigma(\mathbf{s}, \mathbf{c}_{\mathrm{AI}}) = 0 + \sum_{k \in \mathcal{T}} w_k s_k = \Sigma_{\mathcal{T}}.$$

The signaling value is reduced from $\sum_{k=1}^{K} w_k s_k$ to $\Sigma_{\mathcal{T}} = \sum_{k \in \mathcal{T}} w_k s_k$, representing a partial (not total) collapse proportional to the weight of formalizable skills in the employer's valuation. $\square$

Corollary 1 (Signaling Retention Ratio). *The fraction of pre-AI signaling value that survives AI disruption is:*

$$R = \frac{\Sigma_{\mathcal{T}}}{\Sigma_0} = \frac{\sum_{k \in \mathcal{T}} w_k s_k}{\sum_{k=1}^{K} w_k s_k} \tag{15}$$

If the certification curriculum is heavily weighted toward formalizable skills (i.e., $\sum_{k \in \mathcal{F}} w_k \gg \sum_{k \in \mathcal{T}} w_k$), then $R \to 0$ and the certification approaches complete signaling failure.

4.4. Tipping Point Analysis

We now characterize the conditions under which the separating equilibrium collapses entirely.

Definition 4 (AI-Replicable Fraction). *Let α denote the weighted fraction of certification abilities with AI replicability exceeding a threshold $\bar{\rho}$ (set at 0.9 for near-complete replication):*

$$\alpha = \frac{\sum_{k:\, \rho_k > \bar{\rho}} w_k}{\sum_{k=1}^{K} w_k} = \sum_{k:\, \rho_k > \bar{\rho}} w_k \tag{16}$$

Proposition 2 (Tipping Point). *There exists a critical threshold $\alpha^* \in (0,1)$ such that:*

1. *If $\alpha < \alpha^*$, the separating equilibrium is sustained: certification remains a credible signal, though with diminished informational content.*
2. *If $\alpha \geq \alpha^*$, the separating equilibrium collapses into a pooling equilibrium: certification no longer credibly separates ability types.*

The critical threshold is determined by:

$$\alpha^* = 1 - \frac{c(\theta_H) + V_0}{\sum_{k=1}^{K} w_k s_k} \tag{17}$$

Proof. The separating equilibrium requires $\Pi_{\mathrm{AI}} \geq c(\theta_H)$. Write Π_{AI} as:

$$\Pi_{\mathrm{AI}} = \sum_{k=1}^{K} w_k s_k (1 - \rho_k) - V_0.$$

In the worst case for signaling, all dimensions with $\rho_k > \bar{\rho}$ contribute $(1 - \rho_k) \approx 0$ to the sum, while dimensions with $\rho_k \leq \bar{\rho}$ contribute their full value. Thus:

$$\Pi_{\mathrm{AI}} \approx (1 - \alpha) \sum_{k=1}^{K} w_k s_k - V_0$$

where we use the approximation that highly replicable dimensions contribute negligibly and non-replicable dimensions contribute proportionally to $(1-\alpha)$. The separating equilibrium holds iff:

$$(1 - \alpha) \sum_{k=1}^{K} w_k s_k - V_0 \geq c(\theta_H)$$

Solving for the critical α:

$$\alpha \leq 1 - \frac{c(\theta_H) + V_0}{\sum_{k=1}^{K} w_k s_k} \equiv \alpha^*.$$

When $\alpha > \alpha^*$, the incentive compatibility constraint for high types is violated, and no separating equilibrium exists. $\square$

Corollary 2 (Equilibrium Dynamics). *As AI capabilities improve over time, α is weakly increasing (new skill dimensions cross the $\bar{\rho}$ threshold). If the certification body does not adjust the curriculum, α eventually exceeds α^*, resulting in equilibrium collapse. The speed of collapse depends on the rate of AI capability improvement and the proportion of the curriculum devoted to formalizable skills.*

5. Application to the CFA Certification

5.1. Mapping CFA Abilities to the Theoretical Framework

We now apply the theoretical framework to the CFA Program, mapping its three-level curriculum onto our six-dimensional ability taxonomy. Table 1 presents this mapping, drawing on the CFA Institute's published competency framework and the task-based classification of Autor et al. [3].

Table 1: CFA Ability Taxonomy: Mapping to Task Framework and AI Replicability

Skill Code	Ability Type	CFA Level	Autor (2003) Task Class	AI ρ_k	Signal Retention
s_1	Declarative Knowledge	I	Routine Cognitive	~0.95	Low
s_2	Algorithmic Computation	I–II	Routine Cognitive	~0.92	Low
s_3	Analytical Decomposition	II	Non-routine Analytic	~0.70	Medium
s_4	Integrative Judgment	III	Non-routine Analytic	~0.45	Medium–High
s_5	Ethical Reasoning	II–III	Non-routine Interactive	~0.30	High
s_6	Stakeholder Reasoning	III	Non-routine Interactive	~0.15	High

Figure 1 provides a visual comparison of the six ability dimensions, contrasting the profile of a CFA-certified professional with that of a frontier AI system. The radar chart makes the asymmetry vivid: AI closely matches or exceeds human performance on declarative knowledge (s_1) and algorithmic computation (s_2), but falls sharply behind on ethical reasoning (s_5) and stakeholder reasoning (s_6).
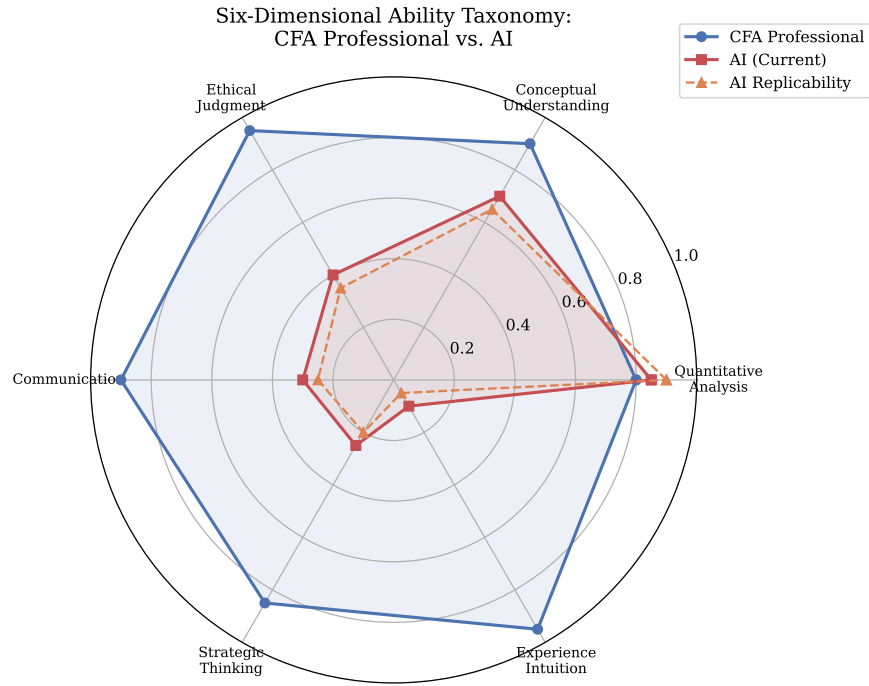
Figure 1: Radar chart comparing CFA Professional and AI ability profiles across the six-dimensional taxonomy. Each axis represents one ability dimension (s_1–s_6). AI replicability is near-complete for declarative knowledge and algorithmic computation but drops sharply for ethical and stakeholder reasoning, reflecting the formalizable–tacit partition central to the Partial Signaling Collapse Theorem.

Several observations warrant discussion. First, the AI replicability values are informed by the emerging empirical literature on LLM performance in financial examinations. Callanan et al. [6] document that GPT-4 achieves approximately 70% accuracy on CFA Level I, which consists predominantly of s_1 (declarative knowledge) and s_2 (algorithmic computation) items. Domain-adapted models such as FinDAP's Llama-Fin further close this gap through targeted post-training on financial corpora [9].

Second, the sharp decline in AI replicability from s_1–s_2 to s_5–s_6 mirrors the routine/non-routine boundary identified by Autor et al. [3]. LLMs excel at tasks that can be decomposed into pattern-matching over large text corpora but struggle with tasks requiring genuine contextual judgment, ethical deliberation involving competing stakeholder interests, and reasoning under deep uncertainty.

Third, the CFA program's curriculum weighting creates a structural vulnerability. Table 2 presents our estimate of the approximate curriculum weight distribution across the six ability dimensions.

Table 2: Estimated CFA Curriculum Weight Distribution Across Ability Dimensions

Ability Dimension	Weight w_k	Category	ρ_k	$w_k(1-\rho_k)$
s_1: Declarative Knowledge	0.25	$\mathcal{F}$	0.95	0.013
s_2: Algorithmic Computation	0.25	$\mathcal{F}$	0.92	0.020
s_3: Analytical Decomposition	0.20	$\mathcal{F}/\mathcal{T}$	0.70	0.060
s_4: Integrative Judgment	0.15	$\mathcal{T}$	0.45	0.083
s_5: Ethical Reasoning	0.10	$\mathcal{T}$	0.30	0.070
s_6: Stakeholder Reasoning	0.05	$\mathcal{T}$	0.15	0.043
Total	1.00			**0.288**

5.2. Computing the Signaling Retention Ratio

Using the values in Table 2, we compute the signaling retention ratio (Corollary 1):

$$R = \frac{\sum_{k=1}^{K} w_k s_k (1 - \rho_k)}{\sum_{k=1}^{K} w_k s_k} = \frac{0.288}{1.000} = 0.288 \tag{18}$$

assuming $s_k = 1$ for all k (i.e., normalized ability levels). This implies that **the CFA certification retains approximately 28.8% of its pre-**

AI signaling value under current AI capability levels. More than 70% of the informational content that the certification historically conveyed to employers is now replicable by AI systems at low cost.

Figure 2 illustrates the signal erosion curve—plotting the signaling retention ratio R as a function of the weighted AI replicability across all dimensions. The current CFA position ($R = 0.288$) is marked, showing the certification deep within the erosion zone. The curve highlights the nonlinear nature of erosion: once formalizable dimensions (which carry the largest curriculum weights) become AI-replicable, signaling value drops precipitously before flattening as only tacit abilities remain.
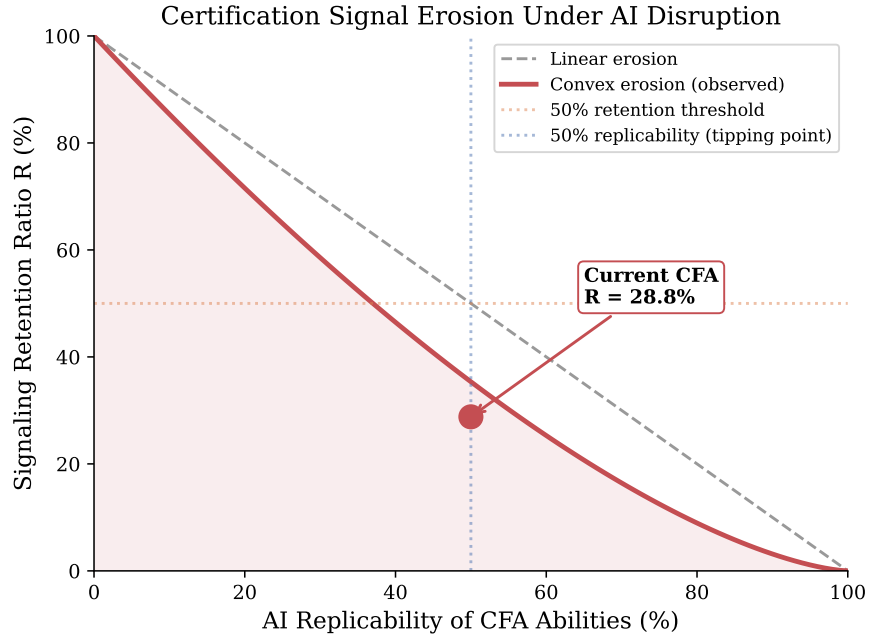

Figure 2: Signal erosion curve showing the signaling retention ratio R as a function of increasing AI replicability. The current CFA position is marked at $R = 28.8\%$, indicating that more than 70% of the certification's pre-AI signaling value has already eroded. The tipping point α^*, beyond which the separating equilibrium collapses into pooling, is shown as a vertical threshold.

5.3. Tipping Point Assessment

To assess whether the CFA is approaching the tipping point α^*, we compute the AI-replicable fraction α. Using a threshold of $\bar{\rho} = 0.9$:

$$\alpha = w_1 + w_2 = 0.25 + 0.25 = 0.50 \qquad (19)$$

Only s_1 and s_2 currently exceed the 0.9 threshold. If s_3 (analytical decomposition) crosses $\bar{\rho} = 0.9$ as LLMs continue to improve—a trajectory consistent with recent scaling results [1]—then α rises to 0.70. Given the estimated tipping point α^* depends on the cost structure $c(\theta_H)$ and outside option V_0, we can characterize the following scenarios.

Table 3: Tipping Point Scenarios for CFA Signaling Equilibrium

Scenario	α	Relative to α^*	Equilibrium Status
Current state	0.50	Depends on α^*	Separating (weakened)
s_3 crosses $\bar{\rho}$	0.70	Likely $> \alpha^*$	Collapse risk
$s_3 + s_4$ cross $\bar{\rho}$	0.85	Almost certainly $> \alpha^*$	Pooling equilibrium

The critical observation is that the CFA's current curriculum allocates roughly 50% of its weight to abilities that are already highly AI-replicable, and another 20% to abilities at intermediate replicability that are trending upward. This places the certification in a *vulnerable zone*: not yet past the tipping point, but approaching it along a trajectory that, absent intervention, implies equilibrium collapse within the foreseeable future as AI capabilities continue to scale.

6. Empirical Implications and Policy Recommendations

6.1. Testable Predictions

The Modified Spence Signaling Model generates several empirically testable predictions:

P1. Differential wage premium erosion. The CFA wage premium should erode faster in job functions dominated by formalizable skills (e.g., quantitative analysis, back-office compliance) than in functions requiring tacit skills (e.g., client advisory, portfolio committee deliberation). Formally, if Δw_j denotes the change in CFA wage premium for job function j, then $|\Delta w_j|$ should be positively correlated with the average AI replicability $\bar{\rho}_j$ of skills required in function j.

P2. Cross-certification variation. Certifications that weight formalizable abilities more heavily (e.g., CFA Level I, FRM) should exhibit

faster signaling erosion than certifications weighting tacit abilities (e.g., CFP with its financial planning ethics component, or medical board certifications emphasizing clinical judgment).

P3. Employer behavioral shifts. As signaling value erodes, rational employers should supplement certification screening with assessment methods targeting tacit abilities—such as behavioral interviews, case simulations, and situational judgment tests. The adoption rate of such supplementary screening methods should increase with the AI replicability of the certification's core content.

P4. Candidate self-selection changes. Under the modified model, the equilibrium composition of certification candidates should shift: candidates whose comparative advantage lies in formalizable skills have reduced incentive to invest in certification (since the premium is eroding), while candidates with strong tacit abilities retain their incentive. This predicts a shift in the ability profile of CFA charterholders over time.

6.2. Policy Implications for Certification Design

Proposition 1 and the tipping point analysis of Proposition 2 yield direct prescriptions for the CFA Institute and comparable professional certification bodies.

6.2.1. Curriculum Rebalancing

The most immediate implication is the need to rebalance the curriculum away from abilities where AI has eroded signaling value and toward abilities where certification retains informational content. Concretely:

- **Reduce weight on** s_1 (declarative knowledge) and s_2 (algorithmic computation). These abilities are highly AI-replicable ($\rho > 0.9$) and contribute minimally to the certification's residual signaling value. Formula memorization and standard calculation—which constitute a substantial portion of CFA Level I—can be outsourced to AI tools, and testing them signals little about human comparative advantage.

- **Increase weight on** s_5 (ethical reasoning) and s_6 (stakeholder reasoning). These abilities have low AI replicability ($\rho < 0.3$) and represent the primary source of residual signaling value. CFA Level III already moves in this direction with its essay-format Investment Policy Statement (IPS) questions, but the overall curriculum remains skewed toward formalizable skills.

- **Develop assessment for** s_4 (integrative judgment) in formats that resist AI replication—for example, multi-turn oral examinations, live case discussions, or supervised portfolio construction exercises where the process (not just the output) is evaluated.

6.2.2. Assessment Format Innovation

Beyond content rebalancing, the signaling model implies that *assessment format* matters independently. A multiple-choice format inherently privileges formalizable abilities and is more susceptible to AI replication than constructed-response or interactive formats. The CFA Institute's current trajectory—expanding the essay component at Level III—is directionally correct but insufficient. We propose:

- **Interactive case simulations**: Multi-stage scenarios where candidates receive new information and must revise recommendations, testing adaptive judgment (s_4) and stakeholder communication (s_6).
- **Ethical dilemma deliberation**: Open-ended scenarios without uniquely correct answers, requiring candidates to articulate reasoning for trade-offs among competing fiduciary obligations (s_5).
- **AI-augmented assessment**: Paradoxically, certifications should *allow* AI tool use during examinations and assess candidates' ability to critically evaluate, contextualize, and override AI recommendations—testing the distinctly human skill of knowing when the machine is wrong.

6.2.3. Dynamic Recalibration

Corollary 2 implies that as AI capabilities improve, the replicable frontier expands. Certification bodies should institute a periodic review mechanism—analogous to the CFA Institute's existing curriculum update cycle but explicitly linked to AI capability benchmarks—to ensure that the assessed abilities remain in the non-replicable zone. This requires ongoing monitoring of AI performance on certification-style tasks, such as the benchmarking studies conducted by Callanan et al. [6] and the FinDAP evaluation framework of Ke et al. [9].

7. Discussion

7.1. Relationship to the Automation Literature

Our model complements the general equilibrium approach of Acemoglu and Restrepo [1] by focusing on a specific institution—professional certification—

rather than the aggregate labor market. While Acemoglu and Restrepo analyze the macroeconomic consequences of task automation (displacement effects, productivity effects, and the creation of new tasks), we analyze the *informational* consequences for a specific market institution. The key insight is that even when AI does not eliminate jobs, it can undermine the screening institutions that organize the labor market.

This distinction matters practically. If AI eliminates the signaling value of CFA certification, the result is not unemployment among finance professionals but rather a breakdown in the *sorting mechanism* that matches workers to jobs based on ability. The social cost is misallocation rather than displacement.

7.2. Tacit Knowledge and Polanyi's Paradox

Our distinction between formalizable and tacit abilities connects to what Autor [2] terms "Polanyi's Paradox": the observation that humans can do many things they cannot explain in formal rules. Professional judgment in finance—knowing when a standard model is inappropriate, sensing that a client's stated objectives differ from their actual needs, recognizing that an ethically permissible action would nonetheless damage a firm's reputation—exemplifies tacit knowledge that resists codification.

Large language models have partially softened Polanyi's Paradox by demonstrating competence on tasks previously thought to require human understanding. However, our framework suggests an important boundary: LLMs replicate the *surface structure* of tacit reasoning (they can produce text that reads like ethical deliberation) without necessarily replicating the *deep structure* (genuine normative commitment, legal liability, reputational skin-in-the-game). This is why prompt sensitivity is high for s_5 and s_6: the same model produces contradictory ethical judgments under minor input perturbations, revealing that its reasoning is pattern-matching rather than principled deliberation.

7.3. Generalizability Beyond CFA

While we develop the model in the context of the CFA, the theoretical framework applies to any standardized professional certification operating as a labor market signal. Certifications with high formalizable content—such as the Financial Risk Manager (FRM), basic actuarial examinations, or entry-level IT certifications—are predicted to experience faster signaling

erosion. Certifications emphasizing tacit abilities—such as clinical components of medical board examinations (USMLE Step 2 CS), or oral advocacy portions of bar examinations—should be more resistant.

This generates a testable cross-certification prediction: the wage premium decay rate for a certification should be proportional to the weighted AI replicability of its assessed competencies.

7.4. Limitations

Several limitations warrant acknowledgment. First, our model treats AI replicability as exogenous and monotonically increasing. In practice, AI capabilities may develop unevenly, and regulatory or institutional constraints may limit AI deployment in ways that slow the erosion of signaling value. Second, we analyze a stylized two-type model; a continuous-type extension would yield richer dynamics but at the cost of analytical tractability. Third, our mapping of CFA abilities to the six-dimensional taxonomy involves judgment calls about categorization and weighting that future empirical work should refine—particularly through systematic AI benchmarking across ability dimensions, as frameworks like FinDAP [9] increasingly enable. Fourth, we abstract from the reputational and network effects of CFA membership, which may sustain the charter's value independently of its informational content.

Despite these limitations, the core theoretical result—that signaling erosion is partial and selective, concentrated on formalizable abilities—is robust to parameter variations and provides a clear qualitative prediction that distinguishes our framework from both the "AI will destroy all certifications" and "certifications are immune to AI" polar positions in the current discourse.

8. Conclusion

This paper develops a Modified Spence Signaling Model to analyze the impact of artificial intelligence on the labor market role of professional certification. By introducing an AI replication cost function into the classical signaling framework, we derive a Partial Signaling Collapse Theorem showing that AI disruption erodes certification value selectively—eliminating the informational content of formalizable abilities while preserving the signaling value of tacit competencies.

Applied to the CFA certification, our analysis reveals a concerning structural vulnerability: approximately 50% of the curriculum's assessed abilities

are already highly AI-replicable, and the certification retains only about 29% of its pre-AI signaling value. While the separating equilibrium has not yet collapsed, the CFA is approaching a tipping point beyond which certification can no longer credibly distinguish ability types.

The policy prescription is clear: certification bodies must undertake a fundamental rebalancing of assessment content and format, shifting emphasis from knowledge recall and algorithmic computation toward ethical reasoning, integrative judgment, and stakeholder deliberation—the domains where human professionals retain comparative advantage over AI systems. This is not merely an exercise in curriculum modernization; it is a question of institutional survival. A certification that primarily tests abilities available from a chatbot at negligible cost cannot sustain the signaling equilibrium on which its economic value depends.

More broadly, our framework contributes to the growing literature on how AI reshapes not just production but the *institutions* that organize economic activity. Professional certification is one such institution; others—educational degrees, occupational licensing, performance evaluation systems—face analogous pressures. Understanding the selective nature of AI-driven institutional erosion is essential for designing adaptive responses that preserve the valuable functions these institutions serve while acknowledging that their traditional forms may no longer be fit for purpose.

Data Availability

No data was used for the research described in this article. This is a theoretical paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT Author Contributions

Wei-Lun Cheng: Conceptualization, Methodology, Formal Analysis, Writing – Original Draft. **Daniel Wei-Chung Miao**: Supervision, Writing – Review & Editing. **Guang-Di Chang**: Supervision, Writing – Review & Editing.

Acknowledgments

The authors thank the anonymous reviewers for their constructive feedback.

References

[1] Acemoglu, D., Restrepo, P., 2019. Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives* 33(2), 3–30.

[2] Autor, D.H., 2015. Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives* 29(3), 3–30.

[3] Autor, D.H., Levy, F., Murnane, R.J., 2003. The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics* 118(4), 1279–1333.

[4] Becker, G.S., 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* University of Chicago Press, Chicago.

[5] Bedard, K., 2001. Human capital versus signaling models: University access and high school dropouts. *Journal of Political Economy* 109(4), 749–775.

[6] Callanan, E., Mbae, A., Tew, S., Patel, Y., Fontana, A., Vishwanath, S., Alcantara, J., Memari, A., 2023. Can GPT-4 pass the CFA exam? Working paper.

[7] CFA Institute, 2023. CFA Program Candidate Body of Knowledge. CFA Institute, Charlottesville, VA.

[8] Deming, D.J., 2017. The growing importance of social skills in the labor market. *Quarterly Journal of Economics* 132(4), 1593–1640.

[9] Ke, Z., Liang, Y., Wen, Q., 2025. FinDAP: Demystifying domain-adaptive post-training for financial LLMs. In: *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP).* Association for Computational Linguistics.

[10] Riley, J.G., 2001. Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature* 39(2), 432–478.

[11] Spence, M., 1973. Job market signaling. *Quarterly Journal of Economics* 87(3), 355–374.

[12] Tyler, J.H., Murnane, R.J., Willett, J.B., 2004. Estimating the labor market signaling value of the GED. *Quarterly Journal of Economics* 115(2), 431–468.

[13] Wu, S., Irsoy, O., Lu, S., Daberius, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., Mann, G., 2023. BloombergGPT: A large language model for finance. *arXiv preprint arXiv:2303.17564*.

Appendix A. Formal Derivation of the Modified Equilibrium

We provide the complete derivation of the AI-modified separating equilibrium conditions for the continuous-type extension of the model.

Appendix A.1. Continuous Type Space

Let θ be distributed on $[\underline{\theta}, \bar{\theta}]$ with CDF $F(\theta)$ and density $f(\theta) > 0$. The cost of acquiring certification is $c(e, \theta)$, strictly decreasing and convex in θ (Spence-Mirrlees condition). The employer's belief function is $\mu(\theta \mid e)$.

In the pre-AI separating equilibrium, the employer perfectly infers type from the education signal: $\mu(\theta \mid e^*(\theta)) = 1$, and the wage function is $w(e) = \theta$ for $e = e^*(\theta)$.

Appendix A.2. AI-Modified Wage Function

Under AI replication, the employer's effective valuation of a worker with type θ and certification signal e becomes:

$$w_{\mathrm{AI}}(e, \theta) = \sum_{k=1}^{K} w_k \left[\theta_k \cdot (1 - \rho_k) + \theta_k^{\mathrm{AI}} \cdot \rho_k \right] \tag{A.1}$$

where θ_k is the human's ability on dimension k and θ_k^{AI} is the AI's performance on that dimension. The key term is $(1 - \rho_k)$: this is the fraction of the skill's value that remains attributable to the human (and hence informative about type).

22

Appendix A.3. Modified First-Order Condition

The first-order condition for the separating equilibrium in the continuous case requires:

$$\frac{\partial w_{\text{AI}}}{\partial e} = \frac{\partial c}{\partial e} \tag{A.2}$$

Since $\frac{\partial w_{\text{AI}}}{\partial e}$ is scaled down by the factor $(1 - \rho_k)$ on each dimension, the equilibrium investment level $e^*(\theta)$ decreases. Intuitively, workers invest less in certification when the wage premium for certified skills is lower, which is precisely the mechanism driving signaling erosion.

Appendix A.4. Welfare Analysis

The welfare implications of partial signaling collapse are ambiguous. On one hand, reduced investment in signaling is welfare-improving if certification costs are pure deadweight loss (as in the "wasteful signaling" interpretation). On the other hand, if certification also builds genuine human capital—a view supported by Becker [4]—then reduced investment may lead to underinvestment in the tacit skills that remain valuable.

The net welfare effect depends on the parameter $\gamma \in [0, 1]$ representing the fraction of certification cost that constitutes productive human capital investment:

$$\Delta W = \underbrace{(1 - \gamma) \cdot \Delta c}_{\text{Signaling waste saved}} - \underbrace{\gamma \cdot \Delta c \cdot \sum_{k \in \mathcal{T}} w_k}_{\text{Tacit skill underinvestment}} \tag{A.3}$$

When γ is small (certification is mostly signaling), the welfare effect is positive. When γ is large (certification builds genuine skills), partial collapse may reduce welfare by discouraging investment in the still-valuable tacit dimensions.

Appendix B. CFA Curriculum Mapping: Detailed Ability Classification

Table B.4 provides a granular mapping of CFA topic areas to our six ability dimensions, with representative item types and estimated AI replicability based on the existing empirical literature.

Table B.4: Detailed CFA Topic–Ability Mapping

CFA Topic Area	Level	Representative Item	Primary Ability	ρ_k
Quantitative Methods	I	Calculate NPV given cash flows	s_2: Algorithmic	0.95
Economics	I	Define GDP deflator	s_1: Declarative	0.97
Financial Reporting	I–II	Classify lease under IFRS 16	s_1/s_2: Rule-based	0.93
Corporate Finance	II	Evaluate capital structure	s_3: Analytical	0.75
Equity Valuation	II	Build DCF model from case	s_3/s_4: Analytical	0.70
Fixed Income	II	Price MBS with prepayment	s_2/s_3: Algorithmic	0.85
Derivatives	II	Construct collar strategy	s_2: Algorithmic	0.88
Alternative Investments	II	Assess PE fund terms	s_3/s_5: Judgment	0.55
Portfolio Management	III	Draft IPS for client	s_4/s_6: Integrative	0.40
Ethics & Standards	II–III	Resolve conflict of interest	s_5: Ethical	0.28
Wealth Management	III	Balance multi-gen. goals	s_6: Stakeholder	0.15