SHOULD HUMANS BE IN THE LOOP? HUMAN-AI COLLABORATION PARADOX AND AUTOMATION CLIFFS

Anonymous authors
Paper under double-blind review

ABSTRACT

Human-in-the-loop AI systems are widely viewed as a safe and effective design for high-stakes decision-making, yet human oversight may weaken as AI becomes more capable. We study how strategic human adaptation shapes organizational human–AI collaboration. We develop a game-theoretic framework in which a coordinator allocates tasks across AI-only, human-only, and AI-assisted workflows while human reviewers endogenously adjust oversight effort. We characterize equilibrium coordination structures and extend the analysis to dynamic learning and heterogeneous multi-task settings. We establish a *Collaboration Paradox*: human–AI collaboration may be suboptimal even when it nominally dominates human-only and AI-only operation, because humans strategically reduce oversight effort as AI becomes more reliable. We also identify *Automation Cliffs*, where small improvements in AI capability trigger abrupt transitions between collaboration, independent operation, and near-full automation. Our results show that the organizational impact of AI depends jointly on technical capability and endogenous human oversight.

1 INTRODUCTION

Human-in-the-loop workflows combine AI recommendations with human judgment and oversight. They are increasingly used in healthcare, software engineering, and professional services (Yin et al., 2025; Peng et al., 2023; Dell'Acqua et al., 2023). Yet human involvement does not guarantee better decisions: reviewers may miss incorrect recommendations or provide only superficial oversight (Goddard et al., 2012; Alon-Barkat & Busuioc, 2023). The relevant question is therefore not simply whether humans and AI can outperform either alone, but whether their potential complementarity survives deployment under limited human capacity and adaptive effort.

Automation bias and complacency offer a key mechanism. As AI becomes more reliable, humans may reduce monitoring and defer to its recommendations (Parasuraman & Riley, 1997; Parasuraman & Manzey, 2010; Dell'Acqua, 2022). Such adaptation need not be irrational: verification is costly, and a reviewer may exert only the effort needed to meet an acceptable performance standard. Consequently, improvements in AI capability can weaken oversight, making realized workflow performance differ from nominal task-level evaluations.

Motivated by these observations, this paper asks a fundamental question: *Does keeping humans in the loop actually improve organizational decision-making?* Our answer is conditional: AI-assisted human review is optimal only in some operational regimes. Depending on relative AI and human capability, the stringency of the performance requirement, and available human capacity, the optimal allocation may instead rely more heavily on AI-only or human-only processing.

To address this question, we develop a game-theoretic framework for human–AI coordination in repeated organizational workflows. A coordinator allocates tasks across AI-only processing, human-only processing, and AI-assisted review; human experts choose their engagement in assisted review. The coordinator maximizes expected performance subject to human capacity, while experts minimize review effort subject to a minimum average performance requirement. These are simultaneous policy choices: neither player commits first, and each player's feasible set depends on the other's decision. We therefore model their interaction as a generalized Nash equilibrium. The three modes follow the levels-of-automation perspective and empirical comparisons of human-only, AI-only, and combined human–AI decision making (Parasuraman et al., 2000; Vaccaro et al., 2024). Applications of this framework include healthcare operations and customer service.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Example 1 (Healthcare Operations)**.** *In oncology-trial prescreening, Parikh et al. (2026) compare unreviewed AI chart extraction, manual eligibility-information abstraction, and human verification of AI-extracted information. These correspond to AI-only, human-only, and AI-assisted processing of prescreening information, preceding enrollment decisions. Radiography offers another analogous comparisons among standalone AI findings, unaided interpretation, and interpretation with AI findings (Patel et al., 2019; Yu et al., 2024).*

Example 2 (Customer Service)**.** *For a case-resolution step, Microsoft Dynamics 365 supports autonomous AI resolution, representative handling without the focal AI suggestion, and representative review of suggested resolutions (Microsoft, 2026b). Zendesk, Genesys Cloud CX, and Salesforce Service Cloud similarly offer autonomous AI agents and agent copilot assistance (Zendesk, 2026a;c; Genesys, 2026; Salesforce, 2026).*

We study routing and review decisions under an institutionally determined resource allocation and performance requirement. Deployable human capacity refers to the staff time or equivalent resources available per incoming task, such as chart-review time or radiographic reporting resources (Parikh et al., 2026; The Royal College of Radiologists, 2022). Customer-service capacity profiles likewise impose agent workload limits (Microsoft, 2026a; Zendesk, 2026b).

The workflow also has an established minimum average performance requirement, such as an accuracy target in healthcare (Parikh et al., 2026; Patel et al., 2019) or a resolution-quality target in customer service (Microsoft, 2026c; Zendesk, 2026d). Service-level agreements specify response- and resolution-time commitments, motivating requirements set separately from work allocation, with an internal minimum average resolution-quality target serving this role. The requirement is fixed because revising it entails changes to clinical governance, quality policies, or customer commitments.

Our contributions can be summarized as follows:

- **Strategic human–AI coordination framework:** We develop a generalized Nash equilibrium framework that jointly characterizes task allocation across three operational modes and endogenous human effort in assisted review, with extensions to dynamic AI learning and heterogeneous tasks. To the best of our knowledge, this paper is among the first to explicitly model human experts as adaptive strategic agents in organizational human–AI collaboration.
- **Managerial implications:** We identify a *Collaboration Paradox*: human–AI collaboration may be suboptimal even when collaboration dominates both human-only and AI-only operation in nominal task-level performance and cost. We further identify *Automation Cliffs*, where small improvements in AI capability can trigger abrupt transitions from human-centered workflows to highly automated operational structures.
- **Empirical validation:** Using radiographic diagnosis as a case study, we show that institutional resource constraints and performance requirements fundamentally shape feasible modes of human–AI coordination. More broadly, our results highlight that the organizational impact of AI adoption depends not only on technical capability, but also on the economic and operational infrastructure supporting human oversight.

Related Work Delegation and learning-to-defer research studies how tasks should be allocated between human experts and algorithms to optimize predictive accuracy or utility (Athey et al., 2020; Mozannar & Sontag, 2020; Wilder et al., 2021; Mozannar et al., 2023). A central focus is learning routing policies from given or estimated performance characteristics, often treating human performance as fixed. We introduce AI-assisted review as a distinct mode in which the human observes and may override an AI recommendation. Review effort responds endogenously to the routing policy and performance requirement, so allocation affects both decision authority and the intensity of human engagement. Consequently, the performance of an assisted workflow depends on its equilibrium allocation rather than solely on the capabilities of its participants.

A second stream studies AI assistance by designing recommendations, feature highlights, or candidate sets to improve human decisions within a specified interaction structure (Li & Zhu, 2024; Noti et al., 2025; Bastani et al., 2026). Our framework makes the allocation across interaction modes an organizational choice: the coordinator selects the shares of AI-only, human-only, and assisted tasks subject to a human-capacity constraint. Human experts choose their engagement in assisted review subject to an aggregate performance requirement. These decisions are mutually dependent because routing determines the review burden, while engagement determines realized performance

and capacity consumption. The resulting equilibrium jointly determines how frequently assistance is used and how intensively humans review it.

A complementary literature examines human supervision and feedback loops in high-stakes AI systems (Ibrahim et al., 2021; De Véricourt & Gurkan, 2023; Zhou et al., 2025; Verma et al., 2025; Balakrishnan et al., 2026). These works highlight how human feedback affects reliability and learning. We formalize this connection in a dynamic environment in which routing decisions determine the quantity and composition of supervision and thereby the law of motion of AI quality.

More broadly, interactive human–algorithm systems generate strategic and dynamic responses (Delarue et al., 2025; Fu et al., 2025; Castro et al., 2025). Closely related work by McLaughlin & Spiess (2024) adopts a principal–agent perspective in which human behavior adapts to algorithmic policies. Our framework shares this focus on strategic adaptation but differs in two respects. We allow randomized triage across three integration modes rather than a binary delegation choice, and we embed this allocation problem in a multi-stage learning environment in which current routing decisions affect future AI performance.

In sum, these literatures motivate the need for models that integrate institutional delegation under capacity constraints, endogenous human engagement under performance standards, and dynamic learning. Our framework provides such an integrated and tractable workflow-level perspective.

2 Core Model: Static Coordination

We begin with a static setting in which the AI system does not improve over time, the coordination policy remains fixed, and all tasks are homogeneous. Consider a setting in which tasks of a single type arrive over time as independent draws from an underlying population. The coordinator chooses a randomized coordination policy $p = (p_A, p_H, p_R)$, where p_m denotes the probability of assigning a task to modality $m \in \{A, H, R\}$ and $p_A + p_H + p_R = 1$. The three modalities correspond to full automation by the algorithm, full human decision making without algorithmic input, and assisted decision making in which the AI issues a recommendation that the human may review. The probabilities p_m represent allocation shares within a homogeneous task class, implemented without task-specific difficulty or confidence signals. Appendix C extends the model to heterogeneous tasks.

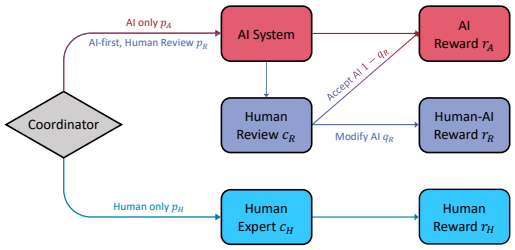


Figure 1: Illustration of human and coordinator policies.

As illustrated by Fig. 1, we now describe each modality in detail:

1. *Full automation (A).* When the coordinator delegates a task to the AI alone, the task is executed automatically and yields expected reward r_A. Without loss of generality, we do not include a per-task algorithmic cost; in many operational settings marginal inference costs are negligible relative to constrained human attention.
2. *Human-only decision making (H).* When the coordinator routes a task to the human without algorithmic input, the expected reward is r_H and the human incurs effort cost c_H per task. The cost c_H captures time, cognitive load, and workflow disruption associated with producing an output without algorithmic assistance.
3. *Assisted decision making (R).* When the coordinator routes a task to assisted decision making, the AI produces a recommendation that is shown to the human. Conditional on receiving a recommendation, the human chooses whether to engage by verifying, editing, or overriding the suggestion. We denote by $q_R \in [0, 1]$ the probability that the human engages when a recommendation is presented. If the human engages, the realized reward is r_R and the human incurs effort cost c_R. If the human does not engage, the recommendation is accepted and the realized reward is r_A. When assisted decision making is used with probability p_R, the expected reward contribution from that modality is $p_R\big(r_R q_R + r_A(1 - q_R)\big)$ and the expected human effort consumed by that modality is $p_R q_R c_R$. Assisted decision making (R) can begin with an AI-generated recom-

mendation (A-H) or with a human requesting AI assistance (H-A). Mode R represents either workflow under the same reward and effort structure for optional human review.

Each of these three modes is well studied in the human–AI collaboration literature. Relative to standard delegation models, we assume that when assisted decision making is used, the coordinator can decide whether an AI recommendation is provided, but the intensity of review is chosen by the human. We represent this discretion by an engagement probability q_R. The core game is simultaneous at the policy level: p and q_R are mutual best responses under coupled constraints, so the term "routing" describes implementation of the equilibrium policy rather than a Stackelberg commitment by the coordinator. We model the human as choosing the least costly level of engagement that satisfies an institutional performance requirement. This captures environments in which review effort is personally costly and not directly contractible, whereas performance is evaluated at the level of aggregate outcomes (Holmström, 1979; Sherstyuk, 2000). Accordingly, the human solves

$$\begin{aligned} \min_{q_R \in [0,1]} \quad & c_R p_R q_R \\ \text{s.t.} \quad & r_H p_H + r_A p_A + \big(r_R q_R + r_A(1-q_R)\big) p_R \geq \delta. \end{aligned} \tag{1}$$

Here, δ denotes the workflow's required performance standard, namely the minimum average level of performance that must be achieved for the workflow to remain acceptable in operation. Higher values of δ correspond to workflows with lower tolerance for underperformance. We interpret δ as a required performance standard that may be externally imposed or internally governed. It is fixed within the equilibrium problem and is not assumed to be costlessly adjustable by the coordinator. For example, in oncology prescreening, δ can specify minimum expected accuracy of eligibility-information abstraction (Parikh et al., 2026). In Appendix B.1, we show that this constrained optimization model can be seen as the solution of a hidden-action incentive model in which the human faces a sufficiently large accountability penalty for aggregate underperformance.

Taking human behavior as given, the coordinator aims to choose a human-AI coordination strategy that maximizes the expected reward subject to a deployable human capacity constraint:

$$\begin{aligned} \max_{p \geq 0} \quad & r_H p_H + r_A p_A + \big(r_R q_R + r_A(1-q_R)\big) p_R \\ \text{s.t.} \quad & c_H p_H + c_R p_R q_R \leq b, \quad p_A + p_H + p_R = 1. \end{aligned} \tag{2}$$

Here, b denotes the deployable human capacity per incoming task during the operating period rather than a staffing or headcount decision. It is measured in the same resource units as c_H and c_R. For example, in oncology prescreening, this is time available per incoming chart (Parikh et al., 2026).

The constraints in (1) and (2) determine which coordination policies are feasible and govern the equilibrium behavior that follows. Because the coordinator's feasible set depends on the human engagement probability through the deployable human-capacity constraint, and the human's feasible set depends on the coordinator's mix of modalities through the performance requirement, the two optimization problems cannot be solved in isolation. We therefore formulate the interaction as a generalized Nash game (Debreu, 1952; Arrow & Debreu, 1954) in which each player optimizes subject to constraint sets that depend on the other player's choice. Proposition 1 provides a simple sufficient condition for the existence of equilibrium. The proof is deferred to Appendix B.2.

Proposition 1 (Existence of Equilibrium for the Core Model)**.** *Suppose that (i) for every $q_R \in [0,1]$ feasible for* (1)*, the coordinator's problem* (2) *has a nonempty feasible set; and (ii) for every p feasible for* (2)*, the human problem* (1) *has a nonempty feasible set. Then the generalized Nash game induced by* (1) *and* (2) *admits at least one equilibrium.*

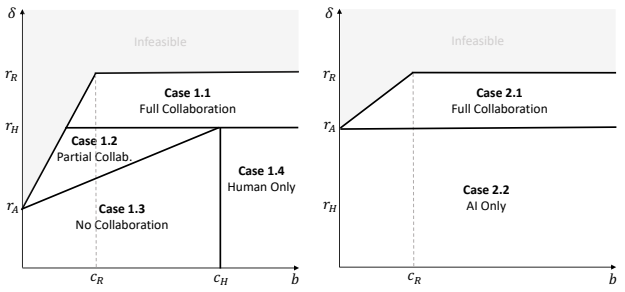
We now characterize equilibrium behavior. To keep the exposition compact and the equilibrium regions interpretable, we impose a weak complementarity condition that reflects common features of assisted decision making.

Assumption 1 (Human-AI Complementarity)**.** $r_R \geq r_H, \quad r_R \geq r_A, \quad c_R \leq c_H.$

Assumption 1 has two parts: (i) When the human *engages* in reviewing an AI recommendation, assisted decision making weakly improves outcomes relative to either agent acting alone. This is consistent with a growing body of evidence showing the benefits of human–AI collaboration across different settings (Vaccaro et al., 2024). Note that this ordering concerns the conditional active-collaboration benchmark r_R; the assisted mode delivers $r_A + (r_R - r_A)q_R$ at engagement level q_R.

(ii) Conditional on engagement, reviewing and editing an AI recommendation is no more costly than producing an output from scratch. This reflects experimental and field evidence that AI assistance often reduces labor input per task in writing and related production workflows (Noy & Zhang, 2023).

Given the assumptions above, we now characterize the equilibrium regions of the core model. Feasibility of the core model is analyzed in Appendix A.1. Detailed derivations and proofs for each equilibrium regime are deferred to Appendix B.4.



(a) Human is better than AI (b) AI is better than human

Figure 2: Equilibrium regions in the core model.

Theorem 1 (Equilibrium Characterization for the Core Model)**.** *Suppose Assumption 1 and the feasibility conditions in Proposition 1 hold. Then the generalized Nash game induced by* (1) *and* (2) *admits an equilibrium. Furthermore, the equilibrium policy $(p^\star, q^\star)$ admits a piecewise characterization over the (δ, b) plane. The geometric structure of these regimes is illustrated in Fig. 2a when humans outperform AI ($r_H > r_A$), and in Fig. 2b when AI weakly outperforms humans ($r_A \geq r_H$). The table in Appendix A.2 summarizes the equilibrium regimes and their associated policies.*

3 Multi-stage and Multi-task Extensions

We focus on a dynamic setting in which the AI system learns over time, generating an endogenous interaction between coordination decisions and future algorithmic performance. Appendix C presents another extension to heterogeneous tasks.

The interaction unfolds over a finite horizon of $T+1$ stages, indexed by $t = 0, 1, \ldots, T$. Let $p_t = (p_A^t, p_H^t, p_R^t) \in \Delta^3$ denote the coordinator's routing policy in stage t, where $p_A^t + p_H^t + p_R^t = 1$ and $p_A^t, p_H^t, p_R^t \geq 0$. Under assisted decision making, the human chooses an engagement level $q_t \in [0, 1]$ in reviewing the recommendation.

The key departure from the static model is that the reward from full automation is no longer fixed. Instead, AI capability improves through exposure to human input. We summarize the accumulated training signal before stage t by a state variable $x_t \geq 0$, with $x_0 = 0$, and write $r_A(x_t)$ for the reward from full automation at that stage.

Assumption 2 (Evolution of AI capability)**.** *The functions $r_A(x_t) : \mathbb{R}_+ \to \mathbb{R}$ and $r_R(x_t) : \mathbb{R}_+ \to \mathbb{R}$ are continuous and nondecreasing in x_t. Moreover, $r_A(x_t) \leq r_R(x_t)$ for all $x_t \geq 0$. The learning state evolves according to*

$$x_{t+1} = x_t + g(x_t, p_t, q_t) = x_t + \lambda p_H^t \mathbf{1}\{r_A(x_t) \leq r_H\} + p_R^t q_t, \tag{3}$$

where g is the learning increment, $\lambda > 0$ measures the informativeness of human-only decisions relative to human engagement with AI recommendations, and $\mathbf{1}\{r_A(x_t) \leq r_H\}$ is the indicator function defined as

$$\mathbf{1}\{r_A(x_t) \leq r_H\} = \begin{cases} 1, & \text{if } r_A(x_t) \leq r_H, \\ 0, & \text{if } r_A(x_t) > r_H. \end{cases}$$

If AI is not as good as humans, the human-only decisions contribute training signal at rate p_H^t, scaled by λ; while the engaged human reviews contribute signal at rate $p_R^t q_t$. In contrast, AI-only decisions do not increase x_t, because the system receives no independent human correction or assessment. This abstraction is closely related to reinforcement learning with human feedback and supervised fine-tuning. The assumption that AI-only decisions do not improve capability is also consistent with recent evidence that recursively training models on their own outputs can degrade rather than improve performance (Shumailov et al., 2024; Alemohammad et al., 2024). When $r_A(x_t) \leq r_H$, $\lambda > 1$ corresponds to a setting in which an unaided human decision provides more training signal than one unit of engaged review under assisted decision making. For notational convenience, we write $r_A^t \equiv r_A(x_t)$ and $r_R^t \equiv r_R(x_t)$ throughout the remainder of the paper.

Given (x_t, p_t, q_t), the expected stage-t reward is weighted average performance across modalities:

$$\ell(x_t, p_t, q_t) = r_A^t p_A^t + r_H p_H^t + \Big(r_A^t + \big(r_R^t - r_A^t\big) q_t\Big) p_R^t. \tag{4}$$

Under assisted decision making, $q_t = 0$ yields the AI-only reward r_A^t, while $q_t = 1$ attains the state-dependent assisted benchmark r_R^t. At each stage, the human chooses the engagement level q_t myopically to minimize effort cost subject to meeting a minimum stage-wise performance requirement δ, as in the core model:

$$q_t \in \arg\min_{q \in [0,1]} c_R p_R^t q \quad \text{s.t.} \quad \ell(x_t, p_t, q) \geq \delta. \tag{5}$$

The human does not internalize the effect of q_t on future AI performance through x_{t+1}.

Anticipating this stage-by-stage best response of the human, the coordinator chooses the routing policy sequence $\{p_t\}_{t=0}^T$ to maximize the total reward over the horizon subject to a per-stage deployable human-capacity constraint:

$$\begin{aligned} \max_{p_0, \cdots, p_T} \quad & \sum_{t=0}^{T} \ell\big(x_t, p_t, q_t^\star(x_t, p_t)\big) \\ \text{s.t.} \quad & c_H p_H^t + c_R p_R^t q_t^\star(x_t, p_t) \leq b, \quad p_t \in \Delta^3, \qquad \forall t \in \{0, 1, \ldots, T\}, \\ & x_{t+1} \text{ follows (3)}, \qquad x_0 = 0. \end{aligned} \tag{6}$$

Here $q_t^\star(x_t, p_t)$ denotes the human's optimal response. Closed-form characterizations of $q_t^\star$ are provided in Appendix F.1.

3.1 Solution to the Multi-stage Problem

Substituting the human best response $q_t^\star(x_t, p_t)$ into the coordinator's objective reduces (6) to a single-agent control problem with state x_t and action p_t. Let

$$\tilde{\ell}(x_t, p_t) \triangleq \ell\big(x_t, p_t, q_t^\star(x_t, p_t)\big), \qquad \tilde{g}(x_t, p_t) \triangleq g\big(x_t, p_t, q_t^\star(x_t, p_t)\big).$$

Restricting attention to policies for which the human expert's problem (5) is feasible, define the feasible action set at state x_t as

$$\mathcal{P}(x_t) \triangleq \left\{p_t \in \Delta^3 : q_t^\star(x_t, p_t) \in [0,1], c_H p_H^t + c_R p_R^t q_t^\star(x_t, p_t) \leq b\right\}.$$

The value function $V_t(x_t)$ is the maximal continuation payoff from stage t onward starting at state x_t. With terminal condition $V_{T+1}(x_{T+1}) \equiv 0$, the Bellman recursion is

$$V_t(x_t) = \max_{p_t \in \mathcal{P}(x_t)} \Big\{\tilde{\ell}(x_t, p_t) + V_{t+1}\big(x_t + \tilde{g}(x_t, p_t)\big)\Big\}, \qquad t = 0, 1, \ldots, T. \tag{7}$$

We solve (7) by backward induction (Bertsekas, 2012; 2017). This formulation highlights the central intertemporal tradeoff: routing affects both current performance and the future learning state through its impact on $x_{t+1} \triangleq x_t + \tilde{g}(x_t, p_t)$. Implementation details are provided in Appendix G.

To facilitate the analysis of the equilibrium states of the dynamic program, we impose the following regularity condition on the evolution of the post-decision learning state.

Assumption 3 (Ordered Learning Dynamics)**.** *We assume the post-decision learning state*

$$x + \tilde{g}\big(x, p^\star(x)\big) = x + \lambda p_H^\star(x) \mathbf{1}\{r_A(x) \leq r_H\} + z\big(x, p_H^\star(x)\big)$$

is nondecreasing in x, where

$$p^\star(x) \triangleq \arg\max_{p \in \mathcal{P}(x)} \tilde{\ell}(x, p), \qquad z(x, p_H^\star(x)) \triangleq \left(\frac{\delta - r_A(x) + \big(r_A(x) - r_H\big) p_H^\star(x)}{r_R(x) - r_A(x)}\right)_+.$$

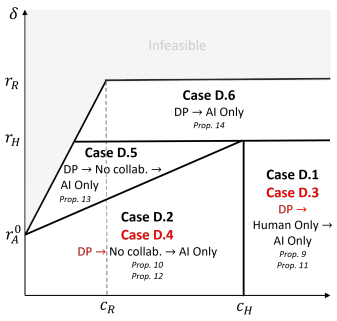
For a fixed state x, $p^\star(x)$ maximizes the current reward after substituting the human's best response. The quantity $z(x, p_H)$ represents the minimum reviewed fraction $p_R q^\star$ required to satisfy the performance requirement δ after fixing the share p_H of tasks routed to human-only decision making. Here $(u)_+ \triangleq \max\{u, 0\}$. Assumption 3 rules out pathological specifications of $r_A(x)$ and $r_R(x)$ in which a system that starts with more accumulated human input could end up with a smaller post-decision learning state under the optimal static routing decision. The condition preserves the interpretation that additional human input monotonically improves the decision system, which we formally record in the following lemma and prove in Appendix F.2.

Lemma 1 (Monotonicity of Value Function)**.** *Suppose Assumptions 2 and 3 hold. Then the value function $V_t(x_t)$ is nondecreasing in x_t for every $t \in \{0, 1, \ldots, T\}$.*

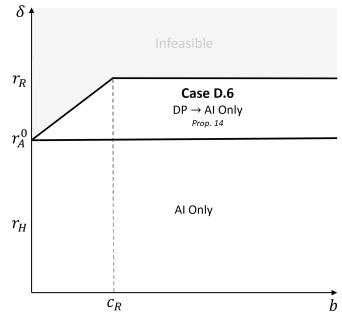
3.2 Analysis of Dynamic Policies

With this regularity condition, we can now analyze the equilibrium in the multi-stage model. First, we will derive the boundary conditions under which the multi-stage problem is feasible in Lemma 2, which is proved in Appendix F.3.

Lemma 2 (Feasibility of the multi-stage model)**.** *Let $\{x_t\}_{t=0}^T$ be the learning-state sequence induced by a feasible policy in* (6)*. Under Assumption 2, the horizon-level feasible region is*



(a) Human is better than AI at $t = 0$



(b) AI is better than human at $t = 0$

Figure 3: Equilibrium regions in the multi-stage model (the dynamic programming (DP) phases and cases shown in red arise only if $\lambda < 1$): (a) $r_H > r_A^0$; (b) $r_A^0 \geq r_H$.

$$\mathcal{I}\big(\{x_t\}_{t=0}^T\big) = \Big\{(\delta, b) :\ \delta \leq M(x_0; b)\Big\} = \Big\{(\delta, b) :\ \delta \leq r_R^0 \text{ and } \delta \leq r_A^0 + \frac{r_R^0 - r_A^0}{c_R} b\Big\}.$$

where:

$$M(x; b) \triangleq \max_{\substack{h,s \geq 0 \\ h+s \leq 1 \\ c_H h + c_R s \leq b}} \{r_A(x)(1-h-s) + r_H h + r_R(x)s\}.$$

The following theorem establishes equilibrium existence and characterizes policy phases as AI performance r_A^t evolves through human input. Derivations are in Appendix F.4.

Theorem 2 (Equilibrium Characterization for the Multi-stage Model)**.** *Suppose Assumptions 2 and 3 hold and the feasibility conditions in Lemma 2 are satisfied. Then the multi-stage model* (6) *admits an equilibrium. Furthermore, the equilibrium policy sequence admits a piecewise threshold characterization over (δ, b) and the evolving AI performance r_A^t, as illustrated in Fig. 3. To be precise, for $\lambda < 1$, define $\kappa_t = \frac{r_H - \lambda r_R^t}{1-\lambda}$, and, when $b < c_H$, define*

$$\theta(\delta, b) = \frac{c_H \delta - b r_H}{c_H - b}, \qquad \rho_t = \begin{cases} \theta(\delta, b), & \lambda \geq 1, \\ \max\{\theta(\delta, b), \kappa_t\}, & \lambda < 1. \end{cases}$$

If $r_A^0 < r_H$ and the corresponding static equilibrium policy is feasible, i.e., either $b \geq c_H$ and $\delta \leq r_H$, or $b < c_H$ and $\delta \leq r_A^0 + \frac{b}{c_H}(r_H - r_A^0)$, then for $\lambda \geq 1$ the equilibrium follows that static allocation whenever $r_A^t < r_H$, whereas for $\lambda < 1$ it may first follow the Bellman solution while $r_A^t < \kappa_t$ and then the corresponding static allocation for $\kappa_t \leq r_A^t < r_H$.

If instead $r_A^0 < r_H$ and $r_A^0 + \frac{b}{c_H}(r_H - r_A^0) < \delta \leq r_H$, the equilibrium follows the Bellman solution while $r_A^t < \rho_t$ and the capacity-constrained mixed allocation for $\rho_t \leq r_A^t < r_H$. If $\delta > r_H$ and $r_A^0 \leq \delta$, the equilibrium follows the Bellman solution while $r_A^t < \delta$. In all cases, once $r_A^t \geq \max\{r_H, \delta\}$, full automation is an equilibrium thereafter.

The complete partition into dynamic equilibrium regions and the exact policies in each region are given in Appendix E.

4 Experiments

We evaluate the human–AI coordination framework through synthetic experiments for the core and multi-stage models and an empirically calibrated illustration of radiographic diagnosis. The synthetic settings are reported in Appendix G. The radiographic illustration applies the core model to chest radiograph diagnosis under Current Procedural Terminology (CPT) code 71046, connecting clinical evidence on standalone and AI-assisted diagnostic performance (Patel et al., 2019; Yu et al.,

2024) with procedure-level reimbursement and cost information organized by CPT codes (American Medical Association, 2026). Appendix H details the empirical calibration. We organize the managerial implications around the *Collaboration Paradox* and *Automation Cliffs* before discussing the additional effects of dynamic AI learning.

Collaboration Paradox Fig. 5a illustrates this implication in the core model: Human–AI collaboration is not universally optimal, even when collaboration achieves both the highest nominal task-level performance ($r_R > \max\{r_H, r_A\}$) and lower marginal human cost ($c_R < c_H$). Instead, collaboration emerges primarily in intermediate operational regimes where AI systems remain imperfect, human expertise is scarce, and institutional requirements are sufficiently stringent that neither pure automation nor pure human processing alone is feasible. Most importantly, the model predicts that increasing the deployable human capacity does not necessarily increase collaboration. In fact, assisted review may disappear from the modeled routing mix as additional human capacity becomes available because humans adapt strategically and minimize costly oversight effort. More details can be found in Appendix A.3.

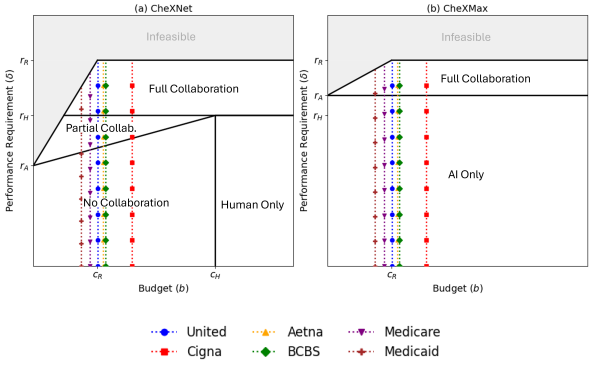
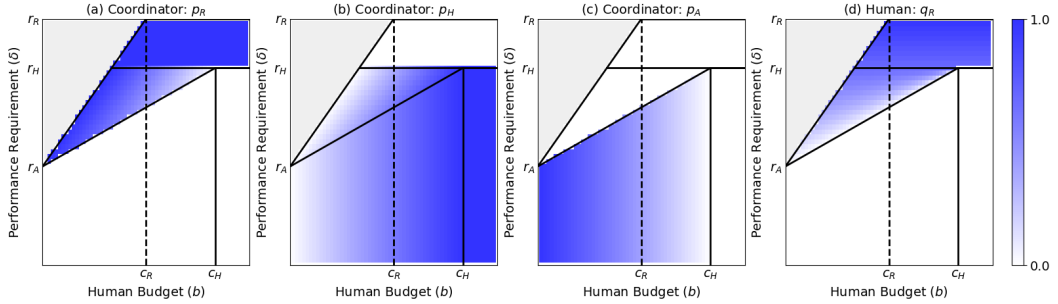
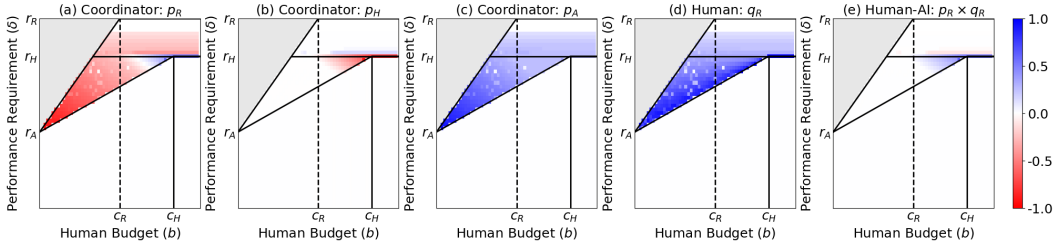


Figure 4: Empirical analysis of radiographic diagnosis.

The multi-stage model preserves this logic even when human input contributes to future AI improvement. When the initial AI baseline r_A^0 lies below the human benchmark r_H and the performance requirement is modest, the optimal policy may forgo assisted review throughout the horizon. Under (3), human-only decisions also provide training information; thus AI learning need not rely on human–AI interaction within the same task. The learning value of human input therefore does not, by itself, make assisted review optimal.



(a) Static policies in the core model when humans outperform AI ($r_H > r_A$).



(b) Difference between the initial policy in the multi-stage model and the static policy in the core model.

Figure 5: Equilibrium policies in the core and multi-stage models.

The radiographic calibration illustrates the same mechanism. In the CheXNet calibration in Fig. 4a, standalone AI is weaker than humans, while the active-review benchmark exceeds both standalone benchmarks r_H and r_A. Reimbursement serves as a proxy for deployable human capacity: when it supports sufficient capacity, providers can meet a given performance requirement by assigning some cases directly to radiologists and the rest to standalone AI. With lower reimbursement, direct human interpretation is more tightly constrained, so providers enter the partial-collaboration region at lower performance requirements.

Automation Cliffs Small changes in the required performance level, deployable human capacity, or AI capability can trigger abrupt transitions between qualitatively different operational modes, including partial collaboration, no collaboration, human-only processing, and full collaboration. A particularly pronounced change occurs when AI reaches the human benchmark: comparing Fig. 2a and Fig. 2b shows that, once $r_A \geq r_H$, the intermediate coordination regimes disappear almost entirely and collapse into AI-only or full-collaboration solutions.

In the multi-stage model, learning can carry the system across these boundaries over time. A policy that initially uses human-only processing or separate human and AI decisions may transition directly to AI-only operation as capability improves. More generally, Theorem 2 establishes that full automation is an equilibrium once $r_A^t \geq \max\{r_H, \delta\}$. Human involvement can consequently support its own eventual withdrawal by improving the AI sufficiently to meet the required standard without oversight. More details can be found in Appendix F.7.

The CheXNet–CheXMax comparison in Fig. 4 illustrates this threshold mechanism under the radiographic calibration. Moving from CheXNet to CheXMax changes the ordering of diagnostic performance from $r_H > r_A$ to $r_A \geq r_H$, largely eliminating the intermediate coordination regimes in Fig. 4(b). The equilibrium then shifts toward either AI-only diagnosis or full collaboration, depending on the required diagnostic standard.

Differences Introduced by Learning Although dynamic AI learning preserves the *Collaboration Paradox* and *Automation Cliffs*, it introduces a future learning value for human involvement, which alters early allocations. Fig. 5b compares the initial multi-stage policy with the static policy in the core model. Panels (a)–(c) show that the shares of tasks assigned to direct human processing (p_H) and to assisted routing (p_R) are typically smaller in the optimal multi-stage policy than in the optimal core policy. This reduction in task volume is offset by higher per-task effort: panel (d) shows a higher engagement probability q_R on routed tasks, and panel (e) shows that the product $p_R q_R$, i.e. the effective reviewed fraction, rises over substantial regions. The coordinator reallocates scarce human effort across modalities to satisfy current-period requirements while increasing the learning signal and accelerating long-run improvement. Consequently, the fraction of tasks receiving engaged review can increase relative to the static benchmark even while the fraction routed to humans falls.

Performance requirements also shape this learning process. Holding the learning state and routing mix fixed, when assisted review is used and the human problem remains feasible, a larger δ weakly raises the minimum engagement needed to satisfy (5). Any resulting increase in the reviewed fraction $p_R^t q_t$ raises the next learning state through (3). This mechanism is illustrated in Fig. 8 in Appendix F.7: among the representative scenarios shown, the green trajectory, which has the highest performance requirement, exhibits the fastest increase in r_A^t. At the same time, a larger δ raises the capability threshold that the AI must reach before full automation becomes feasible.

5 Conclusion

This paper develops a game-theoretic framework for organizational human–AI coordination in which human oversight is endogenous rather than fixed. The analysis shows that collaboration is not universally optimal, even when collaborative decisions outperform both humans and AI in isolation. Instead, the optimal workflow depends on the interaction between AI capability, organizational requirements, and the availability of human expertise.

The analysis reveals automation cliffs, where small changes in AI capability or organizational conditions produce discontinuous changes in workflow design. We further extend the framework to dynamic AI learning environments and heterogeneous task settings, showing how organizations allocate scarce human expertise across tasks and balance current operational performance against future AI improvement.

AI use statement We used generative AI tools to aid and polish the writing of this manuscript, including improving grammar, clarity, readability, and conciseness. The authors reviewed and revised all AI-assisted text to ensure that it preserved the intended meaning and accurately reflected the theoretical models, analyses, and results. We take full responsibility for the final content of this work, including all text produced or revised with the assistance of generative AI.

Ethics statement This work does not involve human subjects, sensitive personal data, or applications that pose direct risks of harm. All datasets used in this paper, along with an anonymized codebase, are provided in the supplementary material, with any potentially sensitive information removed from the real data. We ensured that our methods were applied in a way that does not introduce or amplify unfair bias, discrimination, or privacy risks. The research was conducted in accordance with the ICLR Code of Ethics.

Reproducibility statement We have taken steps to ensure the reproducibility of our results. Experimental details, including dataset descriptions, model architectures, hyperparameters, and implementation of the algorithms, are documented in the main text and the appendix. We have included clear explanations of any assumptions and complete proofs of the theoretical results in both the main paper and the appendix.

REFERENCES

Sina Alemohammad, Josue Casco-Rodriguez, Lorenzo Luzi, Ahmed Imtiaz Humayun, Hossein Babaei, Daniel LeJeune, Ali Siahkoohi, and Richard G. Baraniuk. Self-consuming generative models go MAD. In *Proceedings of the 12th International Conference on Learning Representations*, 2024.

Saar Alon-Barkat and Madalina Busuioc. Human–ai interactions in public sector decision making:"automation bias" and "selective adherence" to algorithmic advice. *Journal of Public Administration Research and Theory*, 33(1):153–169, 2023.

American Medical Association. Current procedural terminology. `https://www.ama-assn.org/practice-management/cpt`, 2026.

Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. Guidelines for human-ai interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019.

Kenneth J Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. In *The Foundations of Price Theory Vol 5*, pp. 289–316. Routledge, London, 1954.

Susan C. Athey, Kevin A. Bryan, and Joshua S. Gans. The allocation of decision authority to human and artificial intelligence. *AEA Papers and Proceedings*, 110:80–84, 2020.

Maya Balakrishnan, Kris Johnson Ferreira, and Jordan Tong. Human-algorithm collaboration with private information: Naïve advice-weighting behavior and mitigation. *Management Science*, 72 (1):265–284, 2026.

Hamsa Bastani, Osbert Bastani, and Wichinpong Park Sinchaisri. Improving human sequential decision making with reinforcement learning. *Management Science*, 72(1):733–755, 2026.

Dimitri Bertsekas. *Dynamic programming and optimal control (4th edition), volume ii*. Athena Scientific, Belmont, MA, 2012.

Dimitri Bertsekas. *Dynamic programming and optimal control (4th edition), volume i*. Athena Scientific, Belmont, MA, 2017.

Dimitri Bertsekas, Angelia Nedic, and Asuman Ozdaglar. *Convex analysis and optimization*, volume 1. Athena Scientific, Belmont, MA, 2003.

Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena scientific, Belmont, MA, 1997.

Francisco Castro, Jian Gao, and Sébastien Martin. Human-ai interactions and societal pitfalls. *arXiv preprint arXiv:2309.10448*, 2025.

Francis De Véricourt and Huseyin Gurkan. Is your machine better than you? you may never know. *Management Science*, 2023.

Gerard Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences*, 38(10):886–893, 1952.

Arthur Delarue, Zhen Lian, and Sebastien Martin. Algorithmic precision and human decision: A study of interactive optimization for school schedules. *Management Science*, 2025.

Fabrizio Dell'Acqua. Falling asleep at the wheel: Human/ai collaboration in a field experiment on hr recruiters. Working paper, 2022. URL https://anacanhoto.com/wp-content/uploads/2024/08/554ee-fallingasleepatthewheel-fabriziodellacqua.pdf.

Fabrizio Dell'Acqua, Edward McFowland III, Ethan Mollick, Hila Lifshitz-Assaf, Katherine C. Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R. Lakhani. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Working Paper 24-013, Harvard Business School, September 2023.

Rabie Adel El Arab and Omayma Abdulaziz Al Moosa. Systematic review of cost effectiveness and budget impact of artificial intelligence in healthcare. *NPJ Digital Medicine*, 8(1):548, 2025.

European Parliament and Council of the European Union. Regulation (eu) 2024/1689 of the european parliament and of the council laying down harmonised rules on artificial intelligence, 2024. Artificial Intelligence Act.

Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer, New York, 2003.

Francisco Facchinei and Jong-Shi Pang. Chapter 12 nash equilibria: the variational approach. *Convex Optimization in Signal Processing and Communications*, pp. 443, 2010.

Runshan Fu, Ginger Zhe Jin, and Meng Liu. Does a human-algorithm feedback loop lead to error propagation? evidence from zillow's zestimate. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pp. 508–508, 2025.

Genesys. Use cases. https://help.mypurecloud.com/usecases/, 2026.

Kate Goddard, Abdul Roudsari, and Jeremy C Wyatt. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1):121–127, 2012.

HCS. Medi-cal rates. https://mcweb.apps.prd.cammis.medi-cal.ca.gov/rates?page=12&tab=rates, 2026.

Joel Hestness, Sharan Narang, Newsha Ardalani, Gregory Diamos, Heewoo Jun, Hassan Kianinejad, Md. Mostofa Ali Patwary, Yang Yang, and Yanqi Zhou. Deep learning scaling is predictable, empirically. *arXiv preprint arXiv:1712.00409*, 2017.

Bengt Holmström. Moral hazard and observability. *The Bell Journal of Economics*, 10(1):74–91, 1979.

Rouba Ibrahim, Song-Hee Kim, and Jordan Tong. Eliciting human judgment for prediction algorithms. *Management Science*, 67(4):2314–2325, 2021.

Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*, 2020.

Michael Lingzhi Li and Shixiang Zhu. Balancing optimality and diversity: Human-centered decision making through generative curation. *arXiv preprint arXiv:2409.11535*, 2024.

Bryce McLaughlin and Jann Spiess. Designing algorithmic recommendations to achieve human-ai complementarity. *arXiv preprint arXiv:2405.01484*, 2024.

MDsave. National average cost for chest x-ray. https://www.mdsave.com/procedures/chest-x-ray/d480face, 2026.

MediBillMD. Ultimate guide to cpt code 71046. https://medibillmd.com/blog/cpt-code-71046/, 2025.

Microsoft. Create and manage capacity profiles. https://learn.microsoft.com/en-us/dynamics365/customer-service/administer/capacity-profiles, 2026a.

Microsoft. Configure Case Management Agent. https://learn.microsoft.com/en-us/dynamics365/customer-service/administer/case-management-global-settings, 2026b.

Microsoft. Overview of service-level agreements. https://learn.microsoft.com/en-us/dynamics365/customer-service/use/overview-service-level-agreements, 2026c.

Hussein Mozannar and David Sontag. Consistent estimators for learning to defer to an expert. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7076–7087, Virtual Event, 2020. PMLR.

Hussein Mozannar, Hunter Lang, Dennis Wei, Prasanna Sattigeri, Subhro Das, and David Sontag. Who should predict? exact algorithms for learning to defer to humans. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 10520–10545, Valencia, Spain, 2023. PMLR.

National Institute of Standards and Technology. Artificial intelligence risk management framework (ai rmf 1.0). Technical report, U.S. Department of Commerce, 2023.

Gali Noti, Kate Donahue, Jon Kleinberg, and Sigal Oren. Ai-assisted decision making with human learning. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pp. 38–39, 2025.

Shakked Noy and Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192, 2023.

Raja Parasuraman and Dietrich H. Manzey. Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3):381–410, 2010.

Raja Parasuraman and Victor Riley. Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2):230–253, 1997.

Raja Parasuraman, Thomas B. Sheridan, and Christopher D. Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30(3):286–297, May 2000. doi: 10.1109/3468.844354.

Ravi B. Parikh, Likhitha Kolla, Elizabeth A. Beothy, William J. Ferrell, Brenda Laventure, Matthew Guido, Anthony Girard, Yang Li, Khaled Essam Mahmoud Dosoky, Karim Tarabishy, Parth S. Patel, Ayana Andalcio, Kristin Maloney, Jose Ulises Mena, Wael Salloum, Jinbo Chen, and Ezekiel J. Emanuel. Human–AI teaming to improve accuracy and efficiency of eligibility criteria prescreening for oncology trials: a randomized evaluation trial using retrospective electronic health records. *Nature Communications*, 17(1):2306, 2026. doi: 10.1038/s41467-026-68873-8. URL https://doi.org/10.1038/s41467-026-68873-8.

Bhavik N Patel, Louis Rosenberg, Gregg Willcox, David Baltaxe, Mimi Lyons, Jeremy Irvin, Pranav Rajpurkar, Timothy Amrhein, Rajan Gupta, Safwan Halabi, et al. Human–machine partnership with artificial intelligence for chest radiograph diagnosis. *NPJ digital medicine*, 2(1):111, 2019.

Payerprice. National average reimbursement for cpt 71046 by major payers. https://payerprice.com/rates/71046-CPT-fee-schedule, 2026.

Sida Peng, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. The impact of ai on developer productivity: Evidence from github copilot. *arXiv preprint arXiv:2302.06590*, 2023.

R Tyrrell Rockafellar and Roger J-B Wets. *Variational Analysis*, volume 317. Springer, Berlin, 2009.

Salesforce. Service Cloud: AI-powered customer service agent console. https://www.salesforce.com/service/cloud/, 2026.

Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, Cambridge, 2014.

Katerina Sherstyuk. Performance standards and incentive pay in agency contracts. *Scandinavian Journal of Economics*, 102(4):725–736, 2000.

Ilia Shumailov, Zakhar Shumaylov, Yiren Zhao, Nicolas Papernot, Ross Anderson, and Yarin Gal. AI models collapse when trained on recursively generated data. *Nature*, 631:755–759, 2024. doi: 10.1038/s41586-024-07566-y.

The Royal College of Radiologists. Radiology reporting figures for service planning 2022. Technical report, The Royal College of Radiologists, London, 2022. https://www.rcr.ac.uk/our-services/all-our-publications/clinical-radiology-publications/radiology-reporting-figures-for-service-planning-2022/.

Michelle Vaccaro, Abdullah Almaatouq, and Thomas Malone. When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, 8:2293–2303, December 2024.

Vladimir N. Vapnik. *Statistical Learning Theory*. Wiley, New York, 1998.

H Satyam Verma, Holly Wiberg, Shixiang Zhu, and Sridhar R Tayur. Data to dose: Efficient synthetic data generation with expert guidance for personalized dosing. *Available at SSRN 5400895*, 2025.

Bryan Wilder, Eric Horvitz, and Ece Kamar. Learning to complement humans. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1526–1533, 2021.

Jiamin Yin, Kee Yuan Ngiam, Sharon Swee-Lin Tan, and Hock Hai Teo. Designing ai-based work processes: How the timing of ai advice affects diagnostic decision making. *Management Science*, 2025.

Feiyang Yu, Alex Moehring, Oishi Banerjee, Tobias Salz, Nikhil Agarwal, and Pranav Rajpurkar. Heterogeneity and predictors of the effects of ai assistance on radiologists. *Nature Medicine*, 30 (3):837–849, 2024.

Zendesk. Creating an AI agent to automatically resolve customer issues. https://support.zendesk.com/hc/en-us/articles/10488757995034-Creating-an-AI-agent-to-automatically-resolve-customer-issues, 2026a.

Zendesk. Creating capacity rules to balance agent workloads. https://support.zendesk.com/hc/en-us/articles/4776409839770-Creating-capacity-rules-to-balance-agent-workloads, 2026b.

Zendesk. About agent copilot. https://support.zendesk.com/hc/en-us/articles/7908817636378-About-agent-copilot, 2026c.

Zendesk. Defining SLA policies. https://support.zendesk.com/hc/en-us/articles/4408829459866-Defining-SLA-policies, 2026d.

Wenbin Zhou, Agni Orfanoudaki, and Shixiang Zhu. Conformalized decision risk assessment. *arXiv preprint arXiv:2505.13243*, 2025.

A Feasibility and Equilibrium Regions of the Core Model

A.1 Feasibility of the Core Model

We first characterize the infeasible region of the parameter space. A pair (δ, b) is infeasible if no combination of delegation and engagement can simultaneously satisfy the coordinator's deployable human capacity and the performance requirement. Because expected reward is bounded above by the best achievable assisted outcome, the most stringent requirement the system can ever satisfy is r_R. Thus, a necessary condition for feasibility is $\delta \leq r_R$. Moreover, even when $\delta \leq r_R$, meeting the requirement may demand a strictly positive level of engagement under assistance, which in turn consumes deployable human capacity. In particular, if the coordinator assigns all tasks to R (the best-performing modality under Assumption 1), the human must engage at least

$$q_{\min}(\delta) = \max\left\{\frac{\delta - r_A}{r_R - r_A}, 0\right\}$$

to achieve average reward δ. This requires human cost $c_R q_{\min}(\delta)$ per task because, for $(p_A, p_H, p_R) = (0, 0, 1)$, $r_R p_R q_{\min}(\delta) + r_A p_R (1 - q_{\min}(\delta)) + r_H p_H + r_A p_A = \delta$. Therefore, a necessary feasibility condition is $b \geq c_R q_{\min}(\delta)$. When either condition $\delta \leq r_R$ or $b \geq c_R q_{\min}(\delta)$ fails, we cannot meet the required performance within the deployable human capacity regardless of how the human responds. Operationally, it corresponds to settings where regulators or administrators demand a level of quality that would require more human review than is available. In such regions, the appropriate response is not a different equilibrium policy but rather a redesign of requirements (lower δ) or resources (increasing b when the capacity constraint is binding).

A.2 Summary of Equilibrium Regions in the Core Model

Case	Region	Condition	$p_A^\star$	$p_H^\star$	$p_R^\star$	$q_R^\star$
1.1 / 2.1	Full collaboration	$\delta > \max\{r_H, r_A\}$ $b \geq c_R q_R^\star$	0	0	1	$\frac{\delta - r_A}{r_R - r_A}$
1.2	Partial collaboration	$r_A + \frac{b}{c_H}(r_H - r_A) < \delta \leq r_H$ $b \geq c_R q_{\min}(\delta)$	0	$\frac{b - c_R q_R^\star}{c_H - c_R q_R^\star}$	$\frac{c_H - b}{c_H - c_R q_R^\star}$	$\frac{(\delta - r_A)c_H + (r_A - r_H)b}{(r_R - r_A)(c_H - b) + (\delta - r_H)c_R}$
1.3	No collaboration	$\delta \leq r_A + \frac{b}{c_H}(r_H - r_A)$ $b < c_H$	$1 - \frac{b}{c_H}$	$\frac{b}{c_H}$	0	0
1.4	Human only	$\delta \leq r_H$ $b \geq c_H$	0	1	0	0
2.2	AI only	$\delta \leq r_A$	1	0	0	0

We assume $r_H > r_A$ under Cases 1.1–1.4, and $r_A \geq r_H$ under Cases 2.1 and 2.2.

A.3 Extra Discussions for the Core Model

As shown in Fig. 6, when AI is at least as capable as humans ($r_A \geq r_H$), the equilibrium structure is relatively straightforward. If the required performance level remains below the AI baseline, the coordinator adopts AI-only operation. Human involvement appears only when the performance requirement approaches or exceeds the AI capability threshold, in which case collaboration becomes necessary to satisfy institutional standards. Consequently, there exist only two economically meaningful regimes: full collaboration (case 2.1) and AI-only (case 2.2).

The more interesting behavior emerges when humans outperform AI ($r_H > r_A$), shown in Fig. 5a. In this setting, the equilibrium depends jointly on the required performance level δ and the deployable human capacity b. A useful way to interpret the figure is to fix δ and study how the optimal policy evolves as the deployable human capacity increases. This perspective partitions the equilibrium structure into three qualitatively distinct regions.

First, when the performance requirement is very stringent ($\delta > r_H$), neither humans nor AI can satisfy the institutional target independently. As a result, collaboration becomes unavoidable. When the deployable human capacity is insufficient, the problem is infeasible because the organization cannot sustain the minimum level of oversight necessary to achieve the required standard. Once b

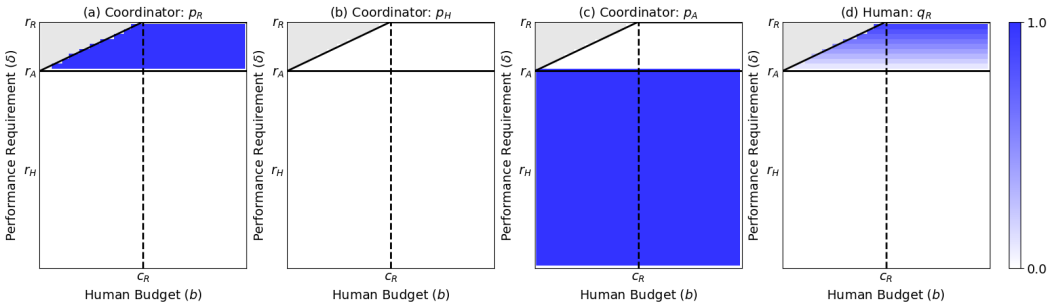


Figure 6: Static policies in the core model when AI is better than humans ($r_A \geq r_H$).

exceeds the feasibility threshold, the equilibrium immediately transitions into full collaboration. Operationally, this region corresponds to high-stakes workflows in which institutions require reliability beyond what either humans or AI can individually achieve.

Second, when the performance requirement satisfies $r_H > \delta > r_A$, the equilibrium transitions from partial collaboration to no collaboration and finally to human only as the deployable human capacity b increases:

$$\text{Partial Collaboration} \;\rightarrow\; \text{No Collaboration} \;\rightarrow\; \text{Human Only},$$

The main observation is that the assisted-review mode can disappear from the routing mix as human capacity increases, even though the active-collaboration benchmark r_R exceeds both standalone benchmarks r_H and r_A. When the deployable human capacity is limited, the coordinator uses collaboration to satisfy the performance requirement while conserving human effort. As b increases, the coordinator can satisfy the same requirement through a separating allocation in which some tasks are assigned entirely to humans and others entirely to AI. Because human review consumes resources that can instead be deployed as primary human decision making, collaboration becomes less attractive as human capacity expands. Consequently, the equilibrium transitions from partial collaboration to no collaboration. Once b exceeds c_H, all tasks can be assigned to humans, yielding the human-only equilibrium. This is a statement about routing for the modeled task stream; it does not imply that the organization should remove AI from other workflows or reduce long-run AI investment. The result therefore shows that the value of collaboration depends not only on the active-collaboration benchmark, but also on how scarce human expertise must be allocated across tasks.

Third, when the performance requirement is modest ($\delta < r_A$), the coordinator can already satisfy the institutional target using AI alone. Consequently, collaboration never emerges in equilibrium. When b is below c_H, the equilibrium lies in the no-collaboration regime, where the coordinator mixes between AI-only and human-only decisions without assisted review. Once b exceeds c_H, the equilibrium transitions directly to human-only operation because humans still provide higher expected performance than AI. Again, "human-only" refers only to routing of this modeled task stream over the modeled operating period; it is not a recommendation that the firm should "abandon AI" elsewhere or alter long-run technology investment.

B Proofs for the Core Model

This section provides detailed proofs underlying Proposition 1, Theorem 1, and the regime-wise closed-form equilibria reported in Section 2. Throughout, we impose Assumption 1 and the feasibility conditions in Proposition 1. We write the coordinator's policy as $p = (p_A, p_H, p_R)$ and the human's engagement as $q \in [0, 1]$ (we suppress the subscript R to simplify the notation).

Recall the coordinator's problem:

$$\begin{aligned} \max_{p \geq 0} \quad & U_C(p, q) := r_H p_H + r_A p_A + [r_A + (r_R - r_A)q]\, p_R \\ \text{s.t.} \quad & c_H p_H + c_R p_R q \leq b, \quad p_A + p_H + p_R = 1, \end{aligned} \tag{8}$$

and the human's problem:

$$\min_{q \in [0,1]} \quad U_H(p, q) := c_R p_R q \quad \text{s.t.} \quad U_C(p, q) \geq \delta. \tag{9}$$

B.1 Incentive Foundation for the Human Engagement Problem

We show that the constrained engagement problem in (1) can be derived from a primitive hidden-action incentive model. The derivation conditions on the maintained human-only execution requirement in Problem Setup; the aggregate penalty governs discretionary review and does not by itself establish that requirement. Fix a routing policy $p = (p_H, p_A, p_R)$ and suppose $p_R > 0$. The human privately chooses an engagement probability $q_R \in [0, 1]$. The human incurs review cost $c_R p_R q_R$, while aggregate workflow performance is observable and evaluated against the required standard δ.

Let $U(p, q_R) = r_H p_H + r_A p_A + \big(r_R q_R + r_A(1 - q_R)\big)p_R$ denote expected aggregate workflow performance. Suppose that violating the performance standard generates a penalty proportional to the shortfall, so that the human solves

$$\min_{q_R \in [0,1]} \quad c_R p_R q_R + \kappa\big(\delta - U(p, q_R)\big)_+, \tag{10}$$

where $(x)_+ = \max\{x, 0\}$ and $\kappa > 0$ measures the strength of the accountability penalty. This penalty can represent heightened monitoring, reputational loss, managerial intervention, or other institutional consequences that arise when aggregate performance falls below the required standard.

We now show that, when the penalty for underperformance is sufficiently large, (10) reduces to the constrained problem in (1). Since $U(p, q_R) = r_H p_H + r_A p_A + r_A p_R + (r_R - r_A)p_R q_R$, aggregate performance is increasing in q_R whenever $r_R > r_A$. Consider any q_R such that $U(p, q_R) < \delta$. In this region, the objective in (10) is $c_R p_R q_R + \kappa\big(\delta - U(p, q_R)\big)$, and its derivative with respect to q_R is $c_R p_R - \kappa(r_R - r_A)p_R$. Thus, if $\kappa(r_R - r_A) > c_R$, the objective is strictly decreasing in q_R whenever the performance standard is violated. In contrast, when $U(p, q_R) \geq \delta$, the penalty term is zero and the objective is simply $c_R p_R q_R$, which is strictly increasing in q_R.

Therefore, if the standard is feasible, namely if $U(p, 1) \geq \delta$, the optimal solution to (10) is the smallest engagement level that satisfies $U(p, q_R) \geq \delta$. Equivalently, the solution is

$$q_R^\star(p) = \left[\frac{\delta - r_H p_H - r_A p_A - r_A p_R}{(r_R - r_A)p_R}\right]_0^1,$$

where $[x]_0^1 = \min\{\max\{x, 0\}, 1\}$. This is exactly the solution to

$$\min_{q_R \in [0,1]} \quad c_R p_R q_R \quad \text{s.t.} \quad r_H p_H + r_A p_A + \big(r_R q_R + r_A(1 - q_R)\big)p_R \geq \delta,$$

which is (1). Hence the constrained engagement problem used in the main text is the reduced form of a hidden-action accountability model in which violating the aggregate performance standard is sufficiently costly relative to the marginal cost of engagement.

B.2 Proof for Feasibility and Existence of Equilibrium

B.2.1 Proof for Proposition 1

The proof is based on Kakutani's fixed point theorem: Let $S \subset \mathbb{R}^n$ be a nonempty, convex compact set. Let $\Phi : S \to S$ be a set-valued mapping such that for each $x \in S$, $\Phi(x)$ is a nonempty closed convex subset of S. If Φ is closed on S, then Φ has a fixed point.

We need to verify the conditions of Kakutani's fixed point theorem. The set S for the generalized Nash equilibrium model consists of the feasible domain of the coordinator (8) and the human (9). Under conditions (i) and (ii) of Proposition 1, both (8) and (9) are linear programs with nonempty feasible sets, it is clear that our set S is nonempty, convex, and compact.

The mapping Φ for our generalized Nash equilibrium model consists of the solution mappings of the coordinator (8) and the human (9). We next show that, for any given point in the set S, the value of the mapping Φ is a nonempty closed convex subset of the set S. Since the linear program (8) has a nonempty bounded constraint set, for any given $q \in [0, 1]$ feasible for the human problem (9), the optimal solution set of the coordinator (8) is nonempty, closed, and convex. Similarly, the linear program (9) has a nonempty bounded constraint set, for any given p feasible for the coordinator's problem (8), the optimal solution set of the human (9) is nonempty, closed, and convex.

It remains to show the mapping Φ for our generalized Nash equilibrium model is closed on set S. To show the solution mappings of the coordinator (8) and the human (9) are closed on their respective domain, we use Proposition 12.1 of Facchinei & Pang (2010). Since the objective functions of the coordinator (8) and the human (9) are continuous, it remains to show that the feasible region for the coordinator (8) is continuous given the human behavior q, and the feasible region for the human (9) is continuous given the coordinator's strategy p. This follows from the well-known fact that a polyhedral graph-convex mapping is Lipschitz continuous on its effective domain (e.g., Example 9.35 of Rockafellar & Wets (2009); Theorem 5.5.8 of Facchinei & Pang (2003)). The proof is thus complete.

B.2.2 Feasible Region

We first formalize the feasibility characterization.

Lemma 3 (Necessary feasibility conditions)**.** *If there exists (p, q) feasible for both* (8) *and* (9)*, then $\delta \le r_R$.*

Proof. Define $\bar{r}_R(q) := r_A + (r_R - r_A)q$. For any p and q, the coordinator's expected reward satisfies

$$U_C(p, q) = r_H p_H + r_A p_A + \bar{r}_R(q) p_R \le \max\{r_H, r_A, \bar{r}_R(q)\} \le r_R,$$

where the last inequality uses $\bar{r}_R(q) \le r_R$ and Assumption 1 (so $r_R \ge r_H, r_A$). Thus $U_C(p, q) \ge \delta$ implies $\delta \le r_R$. □

Next, consider the strongest feasible policy from the coordinator's perspective: allocate all mass to R. Then the human chooses the minimal q satisfying $U_C(p, q) \ge \delta$.

Lemma 4 (Minimal engagement under universal assistance)**.** *Suppose $p_R = 1$. The minimal q satisfying $U_C(p, q) \ge \delta$ is*

$$q_{\min}(\delta) = \begin{cases} \left[\dfrac{\delta - r_A}{r_R - r_A}\right]_+, & \delta \le r_R, \\ \textit{infeasible}, & \delta > r_R. \end{cases}$$

Proof. If $p_R = 1$, then $U_C(p, q) = \bar{r}_R(q) = r_A + (r_R - r_A)q$, which is increasing in q. Solving $r_A + (r_R - r_A)q \ge \delta$ yields the displayed expression and infeasibility when $\delta > r_R$. □

Lemma 5 (Sufficient feasibility condition)**.** *If $\delta \le r_R$ and $b \ge c_R q_{\min}(\delta)$, then there exists (p, q) feasible for both* (8) *and* (9)*.*

Proof. Choose $p_R = 1$ and $q = q_{\min}(\delta)$. Then by Lemma 4, $U_C(p, q) \ge \delta$. The coordinator's deployable human-capacity constraint becomes $c_R q_{\min}(\delta) \le b$, which holds by assumption. Thus (p, q) is jointly feasible. □

Lemmas 3–5 justify the infeasible region discussions in Section A.1: when $\delta > r_R$ or when $\delta > r_A$ but $b < c_R(\delta - r_A)/(r_R - r_A)$, even universal assistance cannot meet the target.

B.2.3 Proof of Theorem 1

Existence follows from Proposition 1. For characterization, Lemma 6 shows that the human best response is always the minimal feasible engagement. For any fixed q, the coordinator problem is a linear program and admits an optimal extreme-point policy (Lemma 8), implying that equilibrium coordination must lie on a finite set of candidate faces, e.g., A only, H only, R only, A–H mix, A–R mix, or H–R mix. For each face, impose the human minimal-feasible response (Lemma 7) and solve the resulting binding equations to obtain closed-form policies; Propositions 2–6 give the explicit solutions for the regimes that arise under Assumption 1. Finally, the region boundaries are obtained by enforcing feasibility ($q \in [0, 1]$, $p \in \Delta$) and comparing the coordinator's objective across candidate extreme points; these algebraic comparisons yield the partitions depicted in Fig. 2.

B.3 Best Responses of Human and Coordinator

B.3.1 Human Best Response: Minimal Engagement

We now solve the human's optimization given a feasible p.

Lemma 6 (Human best response has a threshold form)**.** *Fix p such that the constraint set of* (9) *is nonempty. Then any human best response $q^\star(p)$ is the minimal $q \in [0,1]$ satisfying $U_C(p,q) \geq \delta$. Moreover:*

1. *If $p_R = 0$, then $U_C(p,q)$ does not depend on q, and any feasible q minimizes effort; $q^\star(p) = 0$ is optimal.*
2. *If $p_R > 0$, then $U_C(p,q)$ is affine increasing in q with slope $(r_R - r_A)p_R \geq 0$, and hence the minimal feasible q is unique whenever $r_R > r_A$.*

Proof. The human objective is $U_H(p,q) = c_R p_R q$, which is weakly increasing in q. Therefore the human minimizes effort by choosing the smallest feasible q. If $p_R = 0$, then $U_H(p,q) \equiv 0$ and $U_C(p,q)$ does not depend on q, so $q = 0$ is optimal. If $p_R > 0$, then

$$U_C(p,q) = r_H p_H + r_A p_A + [r_A + (r_R - r_A)q]\, p_R = \underbrace{[r_H p_H + r_A(p_A + p_R)]}_{\text{constant in } q} + (r_R - r_A)p_R\, q,$$

which is affine increasing in q when $r_R > r_A$. Thus, the minimal feasible $q \in [0,1]$ satisfying $U_C(p,q) \geq \delta$ is unique. □

Explicitly, we can derive a closed-form formula for the human (9) as in the lemma below.

Lemma 7 (Closed-form minimal engagement for a given p)**.** *Fix a feasible p with $p_R > 0$ and $r_R > r_A$. Define the baseline reward under the coordinator's policy when the human does not engage: $U_C(p,0) = r_H p_H + r_A(p_A + p_R)$. Then the human best response is*

$$q^\star(p) = \left[\frac{\delta - U_C(p,0)}{(r_R - r_A)p_R}\right]_0^1, \qquad \textit{where } [x]_0^1 := \min\{1, \max\{0, x\}\}.$$

Proof. By Lemma 6, choose the minimal q such that $U_C(p,0) + (r_R - r_A)p_R\, q \geq \delta$. Solving yields $q \geq (\delta - U_C(p,0))/((r_R - r_A)p_R)$, and clipping to $[0,1]$ gives the expression. □

Lemma 7 shows how the coordinator's choice of (p_A, p_H, p_R) determines the human engagement q needed for existence of an equilibrium.

B.3.2 Coordinator Best Response: Extreme-Point Optimality

Fix $q \in [0,1]$. The coordinator's problem (8) is a linear program in p. Hence an optimal solution exists at an extreme point of the feasible polytope.

Lemma 8 (Coordinator optimal policy is an extreme point)**.** *Fix $q \in [0,1]$ such that the feasible set of* (8) *is nonempty. Then there exists an optimal solution $p^\star(q)$ that is an extreme point of the polytope*

$$\mathcal{P}(q) = \Big\{p \geq 0 : p_A + p_H + p_R = 1, c_H p_H + c_R p_R q \leq b\Big\}.$$

Every extreme point of $\mathcal{P}(q)$ has at most two positive components. If two components are positive, the capacity constraint binds.

Proof. The objective of the coordinator (8) is linear in p and the constraints define a nonempty compact polytope. By the fundamental theorem of linear programming, e.g., Theorem 2.7 of Bertsimas & Tsitsiklis (1997), an optimum is attained at an extreme point. A feasible point with three positive components admits a nonzero perturbation preserving their sum and, when binding, capacity; sufficiently small perturbations in both directions remain feasible, so it is not extreme. With two positive components and slack capacity, a small transfer between those components likewise preserves feasibility in both directions. Thus two-mode extreme points must exhaust capacity. □

B.4 Deriving Each Equilibrium Regime

We now derive the closed-form equilibria by combining (i) the human's minimal feasible response (Appendix B.3.1) and (ii) the coordinator's extreme-point optimality (Appendix B.3.2). Because the full partition is summarized in Fig. 2, the goal here is to show how each regime and its policy formulas arise, and to verify equilibrium conditions for each.

B.4.1 AI-only regime

This regime appears as Case 2.2 in Fig. 2b. The proof is similar to Propositions 5 and 6.

Proposition 2 (AI-only equilibrium)**.** *Suppose $\delta \le r_A$ and $r_A \ge r_H$. Then $(p^\star, q^\star)$ with $p_A^\star = 1, p_H^\star = p_R^\star = 0, q^\star = 0$ is an equilibrium.*

B.4.2 Human-only regime

This regime appears as Case 1.4 in Fig. 2a. The proof is similar to the proofs of Propositions 5 and 6.

Proposition 3 (Human-only equilibrium)**.** *Suppose $\delta \le r_H$ and $b \ge c_H$, and additionally $r_H > r_A$. Then $p_H^\star = 1, p_A^\star = p_R^\star = 0, q^\star = 0$ is an equilibrium.*

B.4.3 No-collaboration regime (A–H mix)

Here we derive the no-collaboration equilibrium that appears as Case 1.3 in Fig. 2a. In this regime, $p_R^\star = 0$, $p_H^\star, p_A^\star > 0$, and $q^\star = 0$. The proof is similar to Propositions 5 and 6.

Proposition 4 (No-collaboration equilibrium)**.** *Suppose $\delta \le r_A + b(r_H - r_A)/c_H$, $b < c_H$, and $r_H > r_A$. Then $p_R^\star = 0$, $p_H^\star = b/c_H$, $p_A^\star = 1 - b/c_H$, $q^\star = 0$ is an equilibrium.*

B.4.4 Full collaboration regime (R only)

The full-collaboration equilibrium appears as Case 1.1 in Fig. 2a and Case 1.2 in Fig. 2b. In this regime, $p_R^\star = 1$, $p_A^\star = p_H^\star = 0$, and $q^\star > 0$.

Proposition 5 (Full collaboration equilibrium)**.** *Suppose $\delta > \max\{r_A, r_H\}$, $\delta \le r_R$, and $b \ge c_R(\delta - r_A)/(r_R - r_A)$. Then $p_R^\star = 1$, $p_A^\star = p_H^\star = 0$, $q^\star = (\delta - r_A)/(r_R - r_A)$ is an equilibrium.*

Proof. Fix $p_R = 1$. By Lemma 4, the human best response is $q^\star = (\delta - r_A)/(r_R - r_A)$. Consider the coordinator's best response to $q^\star$. Given $q^\star$, the coordinator's deployable human-capacity constraint becomes $c_R q^\star \le b$, which holds by assumption, so $p_R = 1$ is feasible. Moreover, the achieved reward for the assisted decision making (R) is $r_A + (r_R - r_A)q^\star = r_A + (r_R - r_A) \times (\delta - r_A)/(r_R - r_A) = \delta > \max\{r_A, r_H\}$. Thus, any deviation that places positive mass on A or H would weakly reduce reward and would not reduce required engagement below $q^\star$ in a way that improves the coordinator's objective while maintaining the human constraint at δ. Moreover, back to the perspective of human (9), for fixed $p_R = 1$ the constraint binds at the minimal-feasible $q^\star$, any reduction in p_R must be compensated by increased q, which will raise the human cost and lead to suboptimal solution for the human. Thus, within the region where $p_R = 1$ is feasible, it is also optimal.

To look at it from a different perspective, since $c_R \le c_H$, for given $q^\star$, the coordinator's problem (8) reduces to

$$\max_{p \ge 0} \quad \delta p_R \quad \text{s.t.} \quad c_R q^\star \times p_R \le b, \quad p_R = 1.$$

Naturally, $p_R^\star = 1$ is the optimal policy for the coordinator. □

B.4.5 Partial collaboration regime (H–R mix)

We now derive the partial-collaboration equilibrium that appears as Case 1.2 in Fig. 2a. In this regime, $p_A^\star = 0$, $p_H^\star, p_R^\star > 0$, and $q^\star \in (0, 1]$.

Proposition 6 (Partial collaboration equilibrium). *Suppose $r_H > r_A$, $r_A + b(r_H - r_A)/c_H < \delta \leq r_H$, and $b \geq c_R(\delta - r_A)/(r_R - r_A)$, then the equilibrium policies are*

$$q^\star = \frac{(\delta - r_A)c_H + (r_A - r_H)b}{(r_R - r_A)(c_H - b) + (\delta - r_H)c_R}, \quad p_R^\star = \frac{c_H - b}{c_H - c_R q^\star}, \quad p_H^\star = \frac{b - c_R q^\star}{c_H - c_R q^\star}, \quad p_A^\star = 0. \tag{11}$$

Proof. From $r_H > r_A$ and $r_A + b(r_H - r_A)/c_H < r_H$ we have $b < c_H$. Suppose that the equilibrium policy q for the human (9) satisfies $r_A + (r_R - r_A)q \leq r_H$ and $c_R q \leq b$. Given q, according to Lemma 8, the best response for the coordinator (8) is the point on the upper convex hull of (c_H, r_H), $(c_R q, r_A + (r_R - r_A)q)$, and $(0, r_A)$ with maximum value among all convex combinations whose average human resource use is no larger than b.

Since $r_A \leq r_A + (r_R - r_A)q \leq r_H$ and $c_R q \leq b < c_H$, the best response for the coordinator (8) lies in the convex combination of (c_H, r_H) and $(c_R q, r_A + (r_R - r_A)q)$, subject to the deployable human-capacity constraint:

$$p_R^\star = \frac{c_H - b}{c_H - c_R q}, \qquad p_H^\star = \frac{b - c_R q}{c_H - c_R q}, \qquad p_A^\star = 0.$$

It remains to derive the human's best response $q^\star$ (9), and check that $r_A + (r_R - r_A)q^\star \leq r_H$ and $c_R q^\star \leq b$. Plugging $p_R^\star$, $p_H^\star$, and $p_A^\star$ into the human's problem (9),

$$\min_{q \in [0,1]} \quad \frac{c_R(c_H - b)}{c_H/q - c_R} \quad \text{s.t.} \quad \frac{[(r_R - r_A)(c_H - b) - r_H c_R]q + r_A(c_H - b) + r_H b}{c_H - c_R q} \geq \delta.$$

Since $c_H - c_R q > 0$ (i.e., $c_H/q - c_R > 0$) and $c_H - b > 0$, the human's problem can be further reformulated as:

$$\min_{q \in [0,1]} \quad q \quad \text{s.t.} \quad [(r_R - r_A)(c_H - b) + (\delta - r_H)c_R]q \geq (\delta - r_A)c_H + (r_A - r_H)b.$$

We can show that $(\delta - r_A)c_H + (r_A - r_H)b > 0 \iff \delta > r_A + \frac{r_H - r_A}{c_H} b$.

We can also show that $(r_R - r_A)(c_H - b) + (\delta - r_H)c_R \geq (\delta - r_A)c_H + (r_A - r_H)b \iff \delta \geq \frac{r_R - r_H}{c_R - c_H} b - \frac{c_H r_R - c_R r_H}{c_R - c_H}$.

Notice that the point $(b = c_H, \delta = r_H)$ is on both the two lines $\delta = r_A + \frac{b}{c_H}(r_H - r_A)$ and $\delta = \frac{r_R - r_H}{c_R - c_H} b - \frac{c_H r_R - c_R r_H}{c_R - c_H}$. Given $\delta > r_A + \frac{b}{c_H}(r_H - r_A)$, in order to show that the constraint $\delta \geq \frac{r_R - r_H}{c_R - c_H} b - \frac{c_H r_R - c_R r_H}{c_R - c_H}$ is redundant for $b < c_H$, we only need to compare the two slopes $\frac{r_R - r_H}{c_R - c_H}$ and $\frac{r_H - r_A}{c_H}$:

$$\frac{r_H - r_R}{c_H - c_R} \leq \frac{r_H - r_A}{c_H} \iff c_H(r_H - r_R) \leq c_H(r_H - r_A) - c_R(r_H - r_A)$$
$$\iff \quad c_H(r_A - r_R) \leq -c_R(r_H - r_A) \iff \frac{r_R - r_A}{c_R} \geq \frac{r_H - r_A}{c_H}.$$

The last inequality must hold since $r_R \geq r_H$ and $c_R \leq c_H$ under Assumption 1. Moreover, by setting $b = 0$ for the line $\delta = \frac{r_R - r_H}{c_R - c_H} b - \frac{c_H r_R - c_R r_H}{c_R - c_H}$, then we have $\delta = \frac{c_H r_R - c_R r_H}{c_H - c_R} \geq \frac{c_H r_H - c_R r_H}{c_H - c_R} = r_H$.

As a result, $(r_R - r_A)(c_H - b) + (\delta - r_H)c_R \geq (\delta - r_A)c_H + (r_A - r_H)b > 0$, and the human's best response is

$$q^\star = \frac{(\delta - r_A)c_H + (r_A - r_H)b}{(r_R - r_A)(c_H - b) + (\delta - r_H)c_R} \in (0, 1].$$

It suffices to check that $r_A + (r_R - r_A)q^\star \leq r_H$ and $c_R q^\star \leq b$:

(i) $r_A + (r_R - r_A)q^\star \leq r_H$: Taking the difference between r_H and $r_A + (r_R - r_A)q^\star$, we need to show

$$r_H - [r_A + (r_R - r_A)q^\star] = \frac{(r_H - \delta)[(r_R - r_A)c_H - (r_H - r_A)c_R]}{(r_R - r_A)(c_H - b) + (\delta - r_H)c_R}. \geq 0$$

Since $(r_R - r_A)(c_H - b) + (\delta - r_H)c_R > 0$, it is equivalent to show $\underbrace{(r_H - \delta)}_{\geq 0}[\underbrace{(r_R}_{\geq r_H} - r_A)\underbrace{c_H}_{\geq c_R} - (r_H - r_A)c_R] \geq 0$.

(ii) $c_R q^\star \leq b$: Subtracting $c_R q^\star$ from b, we need to show

$$b - c_R q^\star = \frac{(c_H - b)[b(r_R - r_A) - c_R(\delta - r_A)]}{(r_R - r_A)(c_H - b) + (\delta - r_H)c_R} \geq 0.$$

Since $c_H > b$ and $(r_R - r_A)(c_H - b) + (\delta - r_H)c_R > 0$, it is equivalent to show that

$$b(r_R - r_A) - c_R(\delta - r_A) \geq 0 \quad \Longleftrightarrow \quad b \geq \frac{c_R(\delta - r_A)}{r_R - r_A}.$$

This inequality holds since it is the feasibility assumption we have, as in Lemma 5 and discussed in Section A.1. □

C Model Extension for Heterogeneous Tasks

The core model considers a homogeneous task class. Many operational workflows instead contain tasks that differ in their performance requirements and resource intensity, while drawing on the same pool of human attention. In radiology, for example, low-acuity and high-acuity studies may differ in diagnostic value, clinical risk, and review time, yet the same radiologists may be responsible for both. We extend the static model to this setting by allowing two task classes with type-specific rewards, costs, and performance standards, coupled through a common deployable human capacity.

Task types are indexed by $i \in \{1, 2\}$. Type 1 denotes the low-requirement task and type 2 denotes the high-requirement task. A task is of type i with probability $\lambda_i > 0$, where $\lambda_1 + \lambda_2 = 1$. For each type i, the coordinator chooses a type-specific coordination policy $p_i = (p_{iA}, p_{iH}, p_{iR})$ where $p_{iA} + p_{iH} + p_{iR} = 1$, and the human chooses a type-specific engagement probability $q_i \in [0, 1]$ when a recommendation is presented. The reward and effort primitives are allowed to vary across task types. We denote by r_{iA}, r_{iH}, and r_{iR} the expected rewards from full automation, human-only decision making, and engaged assisted decision making for type i, and by c_{iH} and c_{iR} the corresponding human effort costs under human-only and assisted decision making. The required performance standard for type i is denoted by δ_i.

For each type i, the expected reward induced by (p_i, q_i) is $\ell_i(p_i, q_i) \coloneqq r_{iH}p_{iH} + r_{iA}p_{iA} + \big(r_{iR}q_i + r_{iA}(1 - q_i)\big)p_{iR}$. The human selects the least costly engagement levels that satisfy the type-specific performance requirements:

$$\min_{q_1, q_2 \in [0,1]} \quad \sum_{i=1}^{2} \lambda_i c_{iR} p_{iR} q_i \quad \text{s.t.} \quad \ell_i(p_i, q_i) \geq \delta_i, \quad i = 1, 2. \tag{12}$$

Let v_i denote the utility (weighted by frequency) of type-i task. The coordinator chooses type-specific routing probabilities to maximize expected workflow utility subject to a single deployable human capacity:

$$\begin{aligned} \max_{\{p_i\}_{i=1}^2} \quad & \sum_{i=1}^{2} v_i \ell_i(p_i, q_i) \quad \text{s.t.} \quad \sum_{i=1}^{2} \lambda_i \big(c_{iH}p_{iH} + c_{iR}p_{iR}q_i\big) \leq b, \\ & \qquad\qquad p_{iA} + p_{iH} + p_{iR} = 1, \quad p_i \geq 0, \quad i = 1, 2. \end{aligned} \tag{13}$$

The distinction from the core model is the pooled resource constraint in (13). The coordinator can tailor the routing policy to each task class, and the human can tailor engagement to each task class, but human effort is scarce at the workflow level. Consequently, this shared capacity constraint couples the heterogeneous task classes.

An illustrative example of two heterogeneous task classes is as follows:

Example 3 (Heterogeneous Task Classes)**.** *The two task classes satisfy*

$$\delta_1 \leq \delta_2, \qquad c_{2H} \geq c_{1H}, \qquad c_{2R} \geq c_{1R}, \qquad r_{2A} \geq r_{1A}, \qquad r_{2H} \geq r_{1H}, \qquad r_{2R} \geq r_{1R}.$$

Example 3 captures a common organizational setting in which human decision-makers simultaneously manage both "low-stakes" and "high-stakes" tasks. For example, in hospitals, clinicians may divide time between routine documentation review and clinical assessment. High-stakes tasks generally require a higher performance standard, consume more human effort, and may also generate larger rewards across coordination modalities. In the remainder of this section, we identify the low-stakes and high-stakes tasks with $i = 1$ and $i = 2$, respectively. Note that Example 3 is not an assumption for the analysis.

We impose the following ordering assumption, which extends Assumption 1 to each task class.

Assumption 1A (Type-Specific Human-AI Complementarity)**.** *For each $i \in \{1, 2\}$,*

$$r_{iR} \geq r_{iH}, \qquad r_{iR} \geq r_{iA}, \qquad c_{iR} \leq c_{iH}.$$

Assumption 1A preserves the complementarity structure of the core model within each task class: for both easy and hard tasks, assisted decision-making weakly improves reward over either standalone modality while requiring no more human effort than unaided human decision-making. This ordering assumption mainly serves to facilitate interpretation of the results within the two-task setting; the framework itself naturally extends to more general forms of heterogeneity.

C.1 Feasibility of the Multi-task Extension

The feasibility boundary can be written in closed form. For each type, define

$$q_i^{\min}(\delta_i) = \left(\frac{\delta_i - r_{iA}}{r_{iR} - r_{iA}} \right)_+, \qquad i = 1, 2,$$

where $(x)_+ = \max\{x, 0\}$. This is the minimum engagement probability required for type i if all type-i tasks are assigned to assisted decision making. The minimum human effort required to satisfy both type-specific standards is therefore

$$B_R(\delta_1, \delta_2) = \sum_{i=1}^{2} \lambda_i c_{iR} q_i^{\min}(\delta_i).$$

The following proposition records the resulting feasibility condition. The proof is in Appendix D.2.

Proposition 7 (Feasibility with Heterogeneous Tasks)**.** *Suppose Assumption 1A holds. The two-task static coordination problem is feasible only if $\delta_i \leq r_{iR}$ for $i = 1, 2$ and $B_R(\delta_1, \delta_2) \leq b$. Moreover, these conditions are sufficient.*

The necessity of the first condition follows because r_{iR} is the highest attainable reward for type i under Assumption 1A. The necessity of the second condition follows because $B_R(\delta_1, \delta_2)$ is the least human effort needed to meet the two standards when each type is routed to the most efficient feasible assisted workflow. Sufficiency follows by setting $p_{iR} = 1$ and $q_i = q_i^{\min}(\delta_i)$ for each type with $\delta_i > r_{iA}$, and setting $q_i = 0$ for each type with $\delta_i \leq r_{iA}$.

To distinguish feasibility from the need for collaboration, define the minimum human effort required to satisfy the two standards using only full automation and human-only decision making:

$$B_H(\delta_1, \delta_2) = \sum_{i=1}^{2} \lambda_i c_{iH} \left(\frac{\delta_i - r_{iA}}{r_{iH} - r_{iA}} \right)_+,$$

where the expression is defined for task types with $\delta_i \leq r_{iH}$. If $B_H(\delta_1, \delta_2) \leq b$, both performance requirements can be met without assisted decision making. If instead $B_H(\delta_1, \delta_2) > b$ and $B_R(\delta_1, \delta_2) \leq b$, then assisted decision making is required because the two task classes jointly exhaust the shared deployable human capacity. This case is absent from the single-task model. The driver is not the infeasibility of either task in isolation; rather, the common human resource constraint makes direct human processing too costly across the portfolio of tasks.

The two-task model therefore changes the interpretation of the static equilibrium. In the core model, the pair (δ, b) determines whether a single workflow uses automation, human-only decision making, or assisted decision making. With heterogeneous task classes, the pair (δ_1, δ_2) determines how

scarce human attention is allocated across tasks. Tightening the high-requirement standard δ_2 increases the minimum review effort assigned to type 2, which can change the equilibrium routing of type 1 even when δ_1 and all type-1 primitives remain fixed. The shared deployable human capacity thus creates a cross-task externality: performance requirements imposed on high-acuity tasks can crowd out human oversight for low-acuity tasks. This mechanism provides an additional source of selective collaboration beyond the endogenous engagement channel in the core model.

C.2 Analysis of Heterogeneous Static Policies

We now characterize the equilibrium regions in the heterogeneous static model. The main difference relative to the core model is that the deployable human capacity is shared across task classes. Thus, the analogue of the one-task capacity b is not the aggregate capacity itself, but the amount of human effort allocated to each task class in equilibrium.

For simplicity, we focus on the setting where humans outperform the AI baseline within each task class:

Assumption 5 (Human Advantage Within Task Classes)**.** $r_{iA} < r_{iH} < r_{iR}$, $\forall i \in \{1, 2\}$.

Assumption 5 extends the human-better case in the core model to heterogeneous task classes, preserving the within-task ordering while allowing rewards and costs to vary across task classes.

For each type i, define $\Delta_{iH} = r_{iH} - r_{iA}$, and $\Delta_{iR} = r_{iR} - r_{iA}$. Let $\alpha_i(\delta_i) = ((\delta_i - r_{iA})/\Delta_{iH})_+$ and $\beta_i(\delta_i) = ((\delta_i - r_{iA})/\Delta_{iR})_+$. The quantity $\alpha_i(\delta_i)$ is the minimum share of type-i tasks that must be routed to human-only decision making if assisted review is not used. The quantity $\beta_i(\delta_i)$ is the minimum engagement probability required if all type-i tasks are routed to assisted decision making. Define the corresponding effort thresholds $\underline{z}_i(\delta_i) = c_{iR}\beta_i(\delta_i)$ and $\bar{z}_i(\delta_i) = c_{iH}\alpha_i(\delta_i)$, where $\underline{z}_i(\delta_i)$ is the least type-i human effort needed under full assisted decision making, and $\bar{z}_i(\delta_i)$ is the least type-i human effort needed under human-only routing.

The feasibility frontier from Proposition 7 can be written as $B_R(\delta_1, \delta_2) = \sum_{i=1}^{2} \lambda_i \underline{z}_i(\delta_i)$. When $\delta_i \le r_{iH}$ for both task classes, the no-collaboration frontier is $B_H(\delta_1, \delta_2) = \sum_{i=1}^{2} \lambda_i \bar{z}_i(\delta_i)$. The first frontier determines whether the workflow is feasible with assisted review. The second determines whether both requirements can be met without assisted review. These two frontiers, however, do not fully describe the equilibrium. Between them, the shared capacity must be allocated across task classes, and each task class may be in a different coordination regime.

Let z_i denote the amount of human effort allocated per type-i task: $z_i = c_{iH}p_{iH} + c_{iR}p_{iR}q_i$. The aggregate deployable human capacity constraint is $\sum_{i=1}^{2} \lambda_i z_i \le b$. For type i in the partial-collaboration region, define

$$q_i^P(z_i) = \frac{(\delta_i - r_{iA})c_{iH} - \Delta_{iH} z_i}{\Delta_{iR}(c_{iH} - z_i) + (\delta_i - r_{iH})c_{iR}}. \tag{14}$$

Given $q_i^P(z_i)$, define

$$p_{iH}^P(z_i) = \frac{z_i - c_{iR} q_i^P(z_i)}{c_{iH} - c_{iR} q_i^P(z_i)}, \qquad p_{iR}^P(z_i) = 1 - p_{iH}^P(z_i). \tag{15}$$

These expressions are the type-specific analogues of the partial-collaboration policy in the core model, with the aggregate capacity b replaced by the type-specific allocated effort z_i. The proof is in Appendix D.3.

To state the equilibrium characterization, define the type-i value (reward) from allocating effort z_i after substituting the human's best response (derived in Appendix D.3):

$$W_i(z_i; \delta_i) = \begin{cases} \delta_i, & \underline{z}_i(\delta_i) \le z_i < \bar{z}_i(\delta_i), \quad \delta_i \le r_{iH}, \\ r_{iA} + \dfrac{\Delta_{iH}}{c_{iH}} z_i, & \bar{z}_i(\delta_i) \le z_i \le c_{iH}, \quad \delta_i \le r_{iH}, \\ \delta_i, & z_i = \underline{z}_i(\delta_i), \quad r_{iH} < \delta_i \le r_{iR}. \end{cases}$$

If $\delta_i > r_{iR}$ or $z_i < \underline{z}_i(\delta_i)$, type i is infeasible.

Table 1: Equilibrium regions with two heterogeneous task classes

Case	Region	Condition	$p_{1A}^\star$	$p_{1H}^\star$	$p_{1R}^\star$	$q_1^\star$	$p_{2A}^\star$	$p_{2H}^\star$	$p_{2R}^\star$	$q_2^\star$
3.1	No collaboration	$\mathrm{N}_1, \mathrm{N}_2$	$1 - \frac{z_1^\star}{c_{1H}}$	$\frac{z_1^\star}{c_{1H}}$	0	0	$1 - \frac{z_2^\star}{c_{2H}}$	$\frac{z_2^\star}{c_{2H}}$	0	0
3.2	Type 1 partial collaboration	$\mathrm{P}_1, \mathrm{N}_2$	0	$p_{1H}^P(z_1^\star)$	$p_{1R}^P(z_1^\star)$	$q_1^P(z_1^\star)$	$1 - \frac{z_2^\star}{c_{2H}}$	$\frac{z_2^\star}{c_{2H}}$	0	0
3.3	Type 2 partial collaboration	$\mathrm{N}_1, \mathrm{P}_2$	$1 - \frac{z_1^\star}{c_{1H}}$	$\frac{z_1^\star}{c_{1H}}$	0	0	0	$p_{2H}^P(z_2^\star)$	$p_{2R}^P(z_2^\star)$	$q_2^P(z_2^\star)$
3.4	Dual partial collaboration	$\mathrm{P}_1, \mathrm{P}_2$	0	$p_{1H}^P(z_1^\star)$	$p_{1R}^P(z_1^\star)$	$q_1^P(z_1^\star)$	0	$p_{2H}^P(z_2^\star)$	$p_{2R}^P(z_2^\star)$	$q_2^P(z_2^\star)$
3.5	Type 1 full collaboration	$\mathrm{F}_1, \mathrm{N}_2$	0	0	1	$\beta_1(\delta_1)$	$1 - \frac{z_2^\star}{c_{2H}}$	$\frac{z_2^\star}{c_{2H}}$	0	0
3.6	Type 2 full collaboration	$\mathrm{N}_1, \mathrm{F}_2$	$1 - \frac{z_1^\star}{c_{1H}}$	$\frac{z_1^\star}{c_{1H}}$	0	0	0	0	1	$\beta_2(\delta_2)$
3.7	Type 1 full, type 2 partial	$\mathrm{F}_1, \mathrm{P}_2$	0	0	1	$\beta_1(\delta_1)$	0	$p_{2H}^P(z_2^\star)$	$p_{2R}^P(z_2^\star)$	$q_2^P(z_2^\star)$
3.8	Type 1 partial, type 2 full	$\mathrm{P}_1, \mathrm{F}_2$	0	$p_{1H}^P(z_1^\star)$	$p_{1R}^P(z_1^\star)$	$q_1^P(z_1^\star)$	0	0	1	$\beta_2(\delta_2)$
3.9	Full collaboration	$\mathrm{F}_1, \mathrm{F}_2$	0	0	1	$\beta_1(\delta_1)$	0	0	1	$\beta_2(\delta_2)$

Define the set of feasible capacity allocations for type-i tasks under required performance δ_i as follows:

$$\mathcal{Z}_i(\delta_i) = \begin{cases} [\underline{z}_i(\delta_i), c_{iH}], & \delta_i \le r_{iH}, \\ \{\underline{z}_i(\delta_i)\}, & r_{iH} < \delta_i \le r_{iR}. \end{cases}$$

The equilibrium allocation of the shared capacity is characterized by

$$\max_{z_1, z_2} \quad \sum_{i=1}^{2} v_i W_i(z_i; \delta_i) \qquad \text{s.t.} \quad z_i \in \mathcal{Z}_i(\delta_i), \quad \forall i = 1, 2, \quad \sum_{i=1}^{2} \lambda_i z_i \le b. \tag{16}$$

Let $\mathcal{Z}^\star(b, \delta_1, \delta_2)$ denote the set of solutions to (16). If this set is not singleton, any element of it gives an equilibrium. This set-valuedness is not an artifact of the notation. It reflects the fact that, over part of the partial-collaboration region, additional human effort changes the mix of H and R but does not change realized performance, because the human adjusts engagement to satisfy the requirement at minimum effort.

We use the following type-level regime notation. Type i is in no collaboration, denoted N_i, if $\delta_i \le r_{iH}$ and $z_i^\star \in [\bar{z}_i(\delta_i), c_{iH}]$. Type i is in partial collaboration, denoted P_i, if $r_{iA} < \delta_i \le r_{iH}$ and $z_i^\star \in [\underline{z}_i(\delta_i), \bar{z}_i(\delta_i))$. Type i is in full collaboration, denoted F_i, if $r_{iH} < \delta_i \le r_{iR}$ and $z_i^\star = \underline{z}_i(\delta_i)$. Boundary cases in which two regimes coincide are assigned by closure. The equilibrium characterization with heterogeneous tasks is summarized in Theorem 3 below, and the proof is deferred to Appendix D.4.

Theorem 3 (Equilibrium characterization with heterogeneous tasks)**.** *Suppose Assumptions 1A and 5 hold. If $\delta_i > r_{iR}$ for some i or $B_R(\delta_1, \delta_2) > b$, then no feasible equilibrium satisfies both type-specific performance requirements. Otherwise, for any $(z_1^\star, z_2^\star) \in \mathcal{Z}^\star(b, \delta_1, \delta_2)$, the equilibrium policies are given by Table 1. Conversely, every equilibrium satisfying the capacity-allocation problem* (16) *is represented by one of the regions in Table 1, up to boundary cases.*

C.3 Discussion of Results and Managerial Implications

Table 1 shows that heterogeneity fundamentally changes the structure of the equilibrium. In the homogeneous model, the workflow belongs to a single coordination regime. In the heterogeneous model, each task class may occupy a different regime, so the equilibrium becomes an allocation problem over scarce human attention. The coordinator must determine not only whether to use collaboration, but also where collaboration should be deployed across the task portfolio. Because that attention is shared, changing the requirement or resource intensity of one task class can change the shadow value of capacity and therefore the equilibrium routing of the other task class. This cross-task spillover is a distinct contribution of the heterogeneous model.

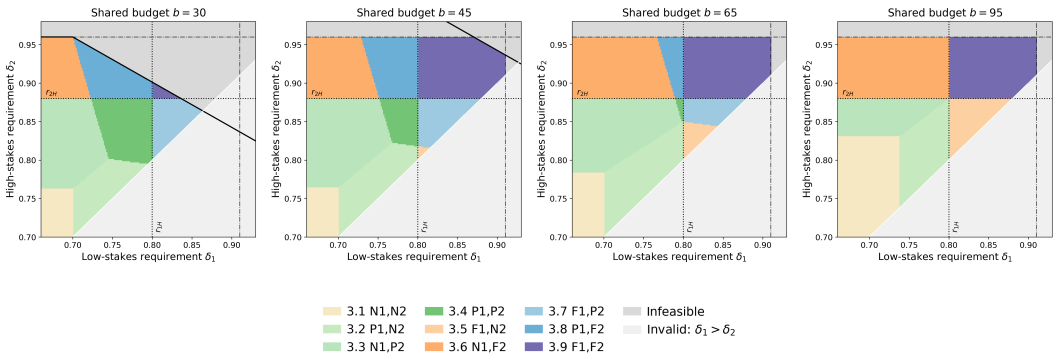


Figure 7: Equilibrium policy under various shared deployable human capacities and performance requirements in the heterogeneous multi-task model (using the setting in Example 3).

Consistency Between the Core and Extended Models The multi-task model preserves the *Collaboration Paradox* in Section 2: the highest active-collaboration benchmark does not guarantee that assisted review is used. Under Assumption 5, r_{iR} exceeds both standalone benchmarks, but the assisted mode delivers $r_{iA} + (r_{iR} - r_{iA})q_i$, which depends on endogenous engagement. For $r_{iA} < \delta_i < r_{iH}$, Table 1 associates feasible allocations below $\bar{z}_i(\delta_i)$ with partial collaboration, allocations at or above this threshold with no collaboration, and $z_i^\star = c_{iH}$ with human-only processing. Thus, additional capacity allocated to a task class can eliminate assisted review despite the higher active-collaboration benchmark. The threshold logic underlying the core model's *Automation Cliffs* also survives: changes in requirements or shared capacity can change which modalities are used for a given class. Because $z_i^\star$ is endogenous, however, neither simultaneous transitions across classes nor monotone growth in each class's collaboration share follows from an increase in b.

Differences Introduced by Heterogeneity A natural intuition is that scarce human effort for AI-assisted review should be concentrated on the high-stakes task as in Example 3. Under Example 3, type 2 has weakly higher requirements, rewards, and human costs than type 1. This logic is also reflected in many discussions of AI deployment, which advocate stronger human oversight for high-risk applications and greater automation for lower-risk tasks (Amershi et al., 2019; National Institute of Standards and Technology, 2023; European Parliament and Council of the European Union, 2024). The equilibrium characterization in Table 1 shows that this intuition is incomplete. Within the model, collaboration depends on each task's requirement relative to its modality benchmarks, together with its human costs, frequency λ_i, utility weight v_i, and the competing demands on shared capacity.

This distinction is illustrated by Case 3.5, the light orange regions in Fig. 7. In this region, type 1 is in full collaboration while type 2 is in no collaboration. This occurs when $r_{1H} < \delta_1 \leq r_{1R}$, $\delta_2 \leq r_{2H}$, and the shared capacity supports $z_1^\star = \underline{z}_1(\delta_1)$ and $z_2^\star \in [\bar{z}_2(\delta_2), c_{2H}]$. The low-stakes task requires assisted review because its requirement exceeds what human-only processing can achieve. By contrast, the high-stakes task can satisfy its requirement without collaboration because its own human-only benchmark remains high. Thus, even though type 2 dominates type 1 in rewards, costs, and requirements, the equilibrium may allocate collaboration to type 1 and not to type 2.

The results show that reserving collaboration for high-stakes tasks can be suboptimal. Collaboration depends on each task's performance requirement relative to its own modality benchmarks and the competing demands on shared human capacity. Consequently, routing decisions should be evaluated jointly across the task portfolio.

The same mechanism explains the partial-collaboration regions (Cases 3.2, 3.3, 3.4, 3.7, 3.8). Type i enters partial collaboration whenever $\underline{z}_i(\delta_i) \leq z_i^\star < \bar{z}_i(\delta_i)$ and $r_{iA} < \delta_i \leq r_{iH}$. In this region, the task could be handled without assisted review if enough human-only capacity were allocated to it. However, because human effort is shared across task classes, assigning that much direct human effort may be too costly at the portfolio level. The coordinator therefore uses assisted review as a capacity-saving substitute for human-only processing. Partial collaboration is not driven purely by

task importance; it is driven by the interaction between type-specific performance requirements and the opportunity cost of assigning shared human capacity to that task class.

The source of this effect is endogenous engagement. Assisted decision making does not automatically deliver the active-collaboration reward r_{iR}; its expected reward is $r_{iA}+(r_{iR}-r_{iA})q_i$, attaining r_{iR} only under full human engagement ($q_i = 1$). Given the routing policy, the human chooses the minimum engagement needed to satisfy the type-specific performance requirement. When the requirement is sufficiently low relative to the human-only benchmark ($\delta_i < r_{iH}$) and enough capacity is allocated to the task class, the coordinator may prefer direct human processing, possibly combined with AI-only processing, because assisted review induces too little engagement to realize its higher performance potential. As the requirement approaches the human-only benchmark ($\delta_i = r_{iH}$), collaboration can become valuable by economizing on scarce human effort. When the requirement is higher than the human-only benchmark ($r_{iH} < \delta_i \le r_{iR}$), collaboration is necessary whenever the workflow is feasible.

Performance Requirements and Cross-Task Spillovers When $\delta_i \le r_{iH}$ for both classes, the frontiers B_R and B_H provide a joint feasibility check. If $B_R \le b < B_H$, assisted review is necessary somewhere in the portfolio. If $B_H \le b$, meeting both standards without assisted review is feasible, but need not maximize the weighted portfolio reward. Tightening δ_2 above r_{2A} raises the minimum capacity reserved for type 2, $\lambda_2 \underline{z}_2(\delta_2)$, and can change the allocation to type 1. For example, a revised allocation with $\underline{z}_1(\delta_1) \le z_1^\star < \bar{z}_1(\delta_1)$ can require partial collaboration for type 1 even though its own standard and technology are unchanged. Both type-specific standards must still be met; if $B_R > b$ or either standard exceeds its active-collaboration benchmark, the portfolio becomes infeasible by Proposition 7.

Consistent with Section 2, a stricter binding performance requirement can increase demand on shared capacity. Externally imposed standards require a portfolio-wide feasibility check; internally governed standards should be assessed jointly with the capacity available to implement them. The equilibrium takes these standards and b as given, so increasing δ_i is not a costless way to induce engagement.

D Proofs for the Heterogeneous Static Model

D.1 Closed-form Solution for the Human's Optimal Response

We first derive the human's best response in the two-task model. Fix a task type i and a routing policy $p_i = (p_{iA}, p_{iH}, p_{iR})$. The type-$i$ performance constraint is $r_{iH}p_{iH} + r_{iA}p_{iA} + \big(r_{iR}q_i + r_{iA}(1-q_i)\big)p_{iR} \ge \delta_i$. Using $p_{iA} + p_{iH} + p_{iR} = 1$, this can be written as $(r_{iR} - r_{iA})p_{iR}q_i + r_{iA} + (r_{iH} - r_{iA})p_{iH} \ge \delta_i$. Thus, whenever $p_{iR} > 0$,

$$q_i \ge \frac{\delta_i - r_{iA} + (r_{iA} - r_{iH})p_{iH}}{(r_{iR} - r_{iA})p_{iR}}.$$

Since the human minimizes $c_{iR}p_{iR}q_i$, the optimal response is the smallest feasible engagement probability. Therefore, on the feasible domain,

$$p_{iR}q_i^\star(p_i) = \left[\frac{\delta_i - r_{iA} + (r_{iA} - r_{iH})p_{iH}}{r_{iR} - r_{iA}}\right]_+. \tag{17}$$

If the right-hand side of (17) exceeds p_{iR}, then the type-i human problem is infeasible under the proposed routing policy. If $p_{iR} = 0$, the human chooses $q_i^\star = 0$ whenever $r_{iA}+(r_{iH}-r_{iA})p_{iH} \ge \delta_i$, and the human problem is infeasible otherwise.

D.2 Proof of Proposition 7: Feasibility

For each type i, let $h_i = p_{iH}$, $s_i = p_{iR}q_i$. The variables h_i and s_i denote, respectively, the fraction of type-i tasks handled by the human alone and the effective reviewed fraction under assisted decision making. The induced reward is $r_{iA}(1 - h_i - s_i) + r_{iH}h_i + r_{iR}s_i = r_{iA} + (r_{iH} - r_{iA})h_i + (r_{iR} - r_{iA})s_i$. The corresponding human effort is $c_{iH}h_i + c_{iR}s_i$.

Under Assumption 1A, $r_{iR} \geq r_{iH}$, $c_{iR} \leq c_{iH}$. Hence assisted review provides weakly larger reward improvement and weakly lower human cost than human-only processing. Therefore, the least costly way to raise type-i performance above the AI baseline is to use assisted review. If $\delta_i > r_{iR}$, even full engagement under assisted decision making cannot satisfy the type-i requirement. Thus $\delta_i \leq r_{iR}$ is necessary.

When $\delta_i \leq r_{iA}$, no human effort is needed for type i. When $\delta_i > r_{iA}$, assigning all type-i tasks to assisted decision making requires engagement $q_i^{\min}(\delta_i) = (\delta_i - r_{iA})/(r_{iR} - r_{iA})$, and consumes type-$i$ effort $c_{iR} q_i^{\min}(\delta_i)$. Thus the minimum aggregate effort needed to satisfy both type-specific standards is

$$B_R(\delta_1, \delta_2) = \sum_{i=1}^{2} \lambda_i c_{iR} \left(\frac{\delta_i - r_{iA}}{r_{iR} - r_{iA}} \right)_+ .$$

Therefore, feasibility requires $B_R(\delta_1, \delta_2) \leq b$.

The same construction proves sufficiency. If $\delta_i \leq r_{iA}$, assign type i to full automation. If $\delta_i > r_{iA}$, assign type i to assisted decision making and set $q_i = q_i^{\min}(\delta_i)$. The resulting aggregate effort is exactly $B_R(\delta_1, \delta_2)$, and each type-specific performance constraint is satisfied. This completes the proof.

D.3 Reduced Coordinator Problem

We next reduce the coordinator's problem after substituting the human's optimal response. Fix type i and write $\Delta_{iH} = r_{iH} - r_{iA}$, $\Delta_{iR} = r_{iR} - r_{iA}$. Using (17), the effective reviewed fraction under the human response is

$$s_i^\star(p_i) = p_{iR} q_i^\star(p_i) = \left[\frac{\delta_i - r_{iA} - \Delta_{iH} p_{iH}}{\Delta_{iR}} \right]_+ .$$

Substituting this expression into the type-i reward gives $\ell_i(p_i) = r_{iA} + \Delta_{iH} p_{iH} + \Delta_{iR} s_i^\star(p_i) = \max\{\delta_i,\, r_{iA} + \Delta_{iH} p_{iH}\}$. The corresponding type-i effort is $z_i(p_i) = c_{iH} p_{iH} + c_{iR} s_i^\star(p_i)$. Thus, after substituting the human response, the coordinator chooses how to allocate human effort across task classes. The only coupling across task classes is the shared deployable human-capacity constraint $\sum_{i=1}^{2} \lambda_i z_i(p_i) \leq b$.

We now characterize the best type-i reward attainable with a given type-level effort z_i. Recall that

$$\underline{z}_i(\delta_i) = c_{iR} \left(\frac{\delta_i - r_{iA}}{\Delta_{iR}} \right)_+ , \qquad \bar{z}_i(\delta_i) = c_{iH} \left(\frac{\delta_i - r_{iA}}{\Delta_{iH}} \right)_+ .$$

First suppose $\delta_i \leq r_{iH}$. If $z_i < \underline{z}_i(\delta_i)$, the type-$i$ requirement is infeasible by Section A.1 and Proposition 7. If $\underline{z}_i(\delta_i) \leq z_i < \bar{z}_i(\delta_i)$, then $p_{iH} \leq c_{iR}\beta_i(\delta_i)/c_{iH} \leq z_i/c_{iH} < \alpha_i(\delta_i)$, so the human-only component cannot satisfy the type-i requirement by itself. The human's performance constraint binds, and the reward is δ_i. This reward is attainable by the partial-collaboration policy in (14)–(15).

If $\bar{z}_i(\delta_i) \leq z_i \leq c_{iH}$, then the coordinator can meet the type-i requirement without assisted review by setting $p_{iH} = z_i/c_{iH}$, $p_{iA} = 1 - z_i/c_{iH}$, $p_{iR} = 0$, $q_i = 0$. This yields reward $r_{iA} + \Delta_{iH} z_i/c_{iH}$. No feasible type-$i$ policy with effort z_i can yield more, because the human response implies $\ell_i(p_i) = \max\{\delta_i,\, r_{iA} + \Delta_{iH} p_{iH}\} \leq \max\{\delta_i,\, r_{iA} + \Delta_{iH} z_i/c_{iH}\}$, and the second term is at least δ_i when $z_i \geq \bar{z}_i(\delta_i)$.

Now suppose $r_{iH} < \delta_i \leq r_{iR}$. Then $r_{iA} + \Delta_{iH} p_{iH} \leq r_{iH} < \delta_i$ for all $p_{iH} \in [0, 1]$. Hence the human's constraint binds for every feasible type-i policy, and the realized reward is δ_i. The least-cost representative is full collaboration: $p_{iR} = 1$, $p_{iH} = p_{iA} = 0$, $q_i = \beta_i(\delta_i)$. This consumes effort $\underline{z}_i(\delta_i)$.

Combining these cases gives the type-level value function $W_i(z_i; \delta_i)$ used in (16).

It remains to verify the partial-collaboration formula. In the partial-collaboration region, the selected representative has $p_{iA} = 0$ and mixes between H and R. Let $h_i = p_{iH}$ and $q_i = q_i^\star$. The effort and performance equations are $c_{iH} h_i + c_{iR} q_i (1 - h_i) = z_i$, and $r_{iA} + \Delta_{iH} h_i + \Delta_{iR} q_i (1 - h_i) = \delta_i$.

Solving these two equations gives

$$q_i = \frac{(\delta_i - r_{iA})c_{iH} - \Delta_{iH} z_i}{\Delta_{iR}(c_{iH} - z_i) + (\delta_i - r_{iH})c_{iR}},$$

which is (14). Substituting this value of q_i into the effort equation gives $p_{iH} = (z_i - c_{iR}q_i)/(c_{iH} - c_{iR}q_i)$, $p_{iR} = 1 - p_{iH}$, which is (15).

D.4 Proof of Theorem 3: Equilibrium Characterization

If $\delta_i > r_{iR}$ for some i or $B_R(\delta_1, \delta_2) > b$, infeasibility follows directly from Proposition 7. Hence no feasible equilibrium can satisfy both type-specific performance constraints.

Suppose instead that the feasibility conditions hold. By Appendix D.3, after substituting the human's best response, the coordinator's problem is equivalent to allocating type-level human effort (z_1, z_2) across the two task classes. The attainable value from assigning effort z_i to type i is $W_i(z_i; \delta_i)$, and the only coupling across task classes is the shared deployable human capacity $\sum_{i=1}^{2} \lambda_i z_i \le b$. Therefore the coordinator's reduced problem is exactly (16).

The feasible set of (16) is nonempty by Proposition 7. It is also compact, because each $\mathcal{Z}_i(\delta_i)$ is compact. The objective $\sum_{i=1}^{2} v_i W_i(z_i; \delta_i)$ is upper semicontinuous on this feasible set: it is piecewise continuous and has only the boundary points corresponding to transitions between partial collaboration and no collaboration, where the two expressions agree. Therefore, according to Proposition 1.2.3 of Bertsekas et al. (2003)), a solution $(z_1^\star, z_2^\star) \in \mathcal{Z}^\star(b, \delta_1, \delta_2)$ exists.

For each solution $(z_1^\star, z_2^\star)$, Appendix D.3 gives the type-level policy that attains $W_i(z_i^\star; \delta_i)$. If $z_i^\star \in [\bar{z}_i(\delta_i), c_{iH}]$ and $\delta_i \le r_{iH}$, the attaining policy is no collaboration: $p_{iA}^\star = 1 - z_i^\star/c_{iH}$, $p_{iH}^\star = z_i^\star/c_{iH}$, $p_{iR}^\star = 0$, $q_i^\star = 0$. If $z_i^\star \in [\underline{z}_i(\delta_i), \bar{z}_i(\delta_i))$ and $r_{iA} < \delta_i \le r_{iH}$, the attaining policy is partial collaboration: $p_{iA}^\star = 0$, $p_{iH}^\star = p_{iH}^P(z_i^\star)$, $p_{iR}^\star = p_{iR}^P(z_i^\star)$, $q_i^\star = q_i^P(z_i^\star)$. If $r_{iH} < \delta_i \le r_{iR}$, the attaining policy is full collaboration: $p_{iA}^\star = p_{iH}^\star = 0$, $p_{iR}^\star = 1$, $q_i^\star = \beta_i(\delta_i)$. Taking the product of the possible type-level regimes for $i = 1, 2$ gives the nine feasible rows in Table 1. The infeasible row follows from Proposition 7. Boundary cases are obtained by closure because the policy expressions coincide at the cutoffs $z_i = \underline{z}_i(\delta_i)$ and $z_i = \bar{z}_i(\delta_i)$. This proves the theorem.

E Solution Existence and Equilibrium Regions of the Multi-stage Model

The next proposition establishes that whenever the multi-stage problem admits a nonempty feasible region, an equilibrium of the coordinator–human interaction exists. Full derivations are deferred to Appendix F.4.

Proposition 8 (Equilibrium characterization for the multi-stage model)**.** *Suppose the feasibility conditions in Lemma 2 hold. Then the multi-stage model* (6) *admits an equilibrium.*

With the solution existence established, we also characterize the equilibrium regions and the associated coordination regimes in the dynamic model. For compactness, we first define several thresholds that govern the transitions. For $\lambda < 1$, define $\kappa_t = \frac{r_H - \lambda r_R^t}{1-\lambda}$, the threshold above which the dynamic value of assisted review is no longer sufficient to displace the corresponding static allocation. Similarly, define $\theta(\delta, b) = \frac{c_H \delta - b r_H}{c_H - b}$, the AI performance level at which the capacity-constrained mixed allocation first satisfies the performance requirement. Finally, define

$$\rho_t = \begin{cases} \theta(\delta, b), & \lambda \ge 1, \\ \max\{\theta(\delta, b), \kappa_t\}, & \lambda < 1. \end{cases}$$

Let $(p_A^B(t), p_H^B(t), p_R^B(t), q^B(t))$ denote any solution of the Bellman equation in (7). We use the following region notation:

$$\mathcal{D}_1 = \{b \geq c_H,\ \delta \leq r_H,\ r_A^0 < r_H,\ \lambda \geq 1\},$$
$$\mathcal{D}_2 = \left\{b < c_H,\ \delta \leq r_A^0 + \frac{b}{c_H}(r_H - r_A^0),\ r_A^0 < r_H,\ \lambda \geq 1\right\},$$
$$\mathcal{D}_3 = \{b \geq c_H,\ \delta \leq r_H,\ r_A^0 < r_H,\ \lambda < 1\},$$
$$\mathcal{D}_4 = \left\{b < c_H,\ \delta \leq r_A^0 + \frac{b}{c_H}(r_H - r_A^0),\ r_A^0 < r_H,\ \lambda < 1\right\},$$
$$\mathcal{D}_5 = \left\{\delta \leq M(x_0; b),\ b < c_H,\ r_A^0 < r_H,\ r_A^0 + \frac{b}{c_H}(r_H - r_A^0) < \delta \leq r_H\right\},$$
$$\mathcal{D}_6 = \{\delta \leq M(x_0; b),\ \delta \geq r_H,\ r_A^0 \leq \delta\}.$$

With this notation, the corresponding dynamic equilibrium policies are summarized in Table 2. The formal propositions and proofs of the dynamic equilibrium regions are contained in Appendices F.5 and F.6. An overview of the dynamic equilibrium structure in the original state space is illustrated in Figure 3. Numerical illustrations of the evolution of the equilibrium policies are provided in Appendix F.7.

Table 2: Summary of dynamic equilibrium regions.

Case	Region	Condition	p_A^t	p_H^t	p_R^t	q_t
D.1	Human only	$\mathcal{D}_1,\ r_A^t < r_H$	0	1	0	0
D.1	Automation	$\mathcal{D}_1,\ r_A^t \geq r_H$	1	0	0	0
D.2	Mixed human and AI	$\mathcal{D}_2,\ r_A^t < r_H$	$1 - \frac{b}{c_H}$	$\frac{b}{c_H}$	0	0
D.2	Automation	$\mathcal{D}_2,\ r_A^t \geq r_H$	1	0	0	0
D.3	Bellman	$\mathcal{D}_3,\ r_A^t < \kappa_t$	$p_A^B(t)$	$p_H^B(t)$	$p_R^B(t)$	$q^B(t)$
D.3	Human only	$\mathcal{D}_3,\ \kappa_t \leq r_A^t < r_H$	0	1	0	0
D.3	Automation	$\mathcal{D}_3,\ r_A^t \geq r_H$	1	0	0	0
D.4	Bellman	$\mathcal{D}_4,\ r_A^t < \kappa_t$	$p_A^B(t)$	$p_H^B(t)$	$p_R^B(t)$	$q^B(t)$
D.4	Mixed human and AI	$\mathcal{D}_4,\ \kappa_t \leq r_A^t < r_H$	$1 - \frac{b}{c_H}$	$\frac{b}{c_H}$	0	0
D.4	Automation	$\mathcal{D}_4,\ r_A^t \geq r_H$	1	0	0	0
D.5	Bellman	$\mathcal{D}_5,\ r_A^t < \rho_t$	$p_A^B(t)$	$p_H^B(t)$	$p_R^B(t)$	$q^B(t)$
D.5	Mixed human and AI	$\mathcal{D}_5,\ \rho_t \leq r_A^t < r_H$	$1 - \frac{b}{c_H}$	$\frac{b}{c_H}$	0	0
D.5	Automation	$\mathcal{D}_5,\ r_A^t \geq r_H$	1	0	0	0
D.6	Bellman	$\mathcal{D}_6,\ r_A^t < \delta$	$p_A^B(t)$	$p_H^B(t)$	$p_R^B(t)$	$q^B(t)$
D.6	Automation	$\mathcal{D}_6,\ r_A^t \geq \delta$	1	0	0	0

Several qualitative patterns emerge from Table 2. When $\lambda \geq 1$, human-only processing generates at least as much learning signal as assisted review. Dynamic and static incentives are therefore aligned in $\mathcal{D}_1$ and $\mathcal{D}_2$, and the coordinator follows the corresponding static equilibrium until AI performance reaches r_H, after which the system transitions to automation.

When $\lambda < 1$, assisted review generates more learning per unit of engaged mass than human-only processing. The coordinator may therefore begin in a Bellman phase that sacrifices current reward to accelerate future AI improvement. Once AI performance reaches the relevant threshold, the value of additional learning diminishes and the policy transitions to the corresponding static allocation before eventually converging to automation.

The performance requirement determines the nature of this transition. In $\mathcal{D}_5$, the mixed allocation is initially infeasible and becomes optimal only after AI performance reaches $\theta(\delta, b)$, which is increasing in δ and decreasing in b. In $\mathcal{D}_6$, the requirement exceeds the human benchmark itself, eliminating human-only processing as a feasible terminal regime. The system therefore remains in the Bellman phase until AI performance reaches δ, after which it transitions directly to automation.

F PROOFS FOR THE MULTI-STAGE MODEL

Throughout this section, we use r_A^t and r_R^t to represent $r_A(x_t)$ and $r_R(x_t)$, respectively, to simplify notation.

F.1 CLOSED-FORM SOLUTION FOR THE HUMAN'S OPTIMAL RESPONSE IN THE MULTI-STAGE MODEL

In this appendix, we derive the closed-form solution for the human's problem (5) at each stage $t \in \{0, 1, 2, \ldots, T\}$. From the constraint $\ell(x_t, p_t, q_t) \geq \delta$, we have $\ell(x_t, p_t, q_t) = r_A^t\, p_A^t + r_H\, p_H^t + \Big(r_A^t + \big(r_R^t - r_A^t\big) q_t\Big)\, p_R^t = \big(r_R^t - r_A^t\big) p_R^t q_t + r_A^t\, p_A^t + r_H\, p_H^t + r_A^t\, p_R^t \geq \delta$.

Since $r_R^t - r_A^t > 0$ by definition, we have $p_R^t q_t \geq (\delta - r_A^t\, p_A^t - r_H\, p_H^t - r_A^t\, p_R^t)/(r_R^t - r_A^t)$. Recall that we also have $q_t \in [0, 1]$. For a given policy p_t of the coordinator in stage t, the human (5) aims to minimize the effort cost, $c_R\, p_R^t q_t$. There are four possible situations.

(i) When $p_R^t = 0$: If $\delta - r_A^t p_A^t - r_H p_H^t \leq 0$, then we have $q_t^\star(x_t, p_t) = 0$. If $\delta - r_A^t p_A^t - r_H p_H^t > 0$, the problem is infeasible.

(ii) When $p_R^t \neq 0$: $q_t \geq (\delta - r_A^t p_A^t - r_H p_H^t - r_A^t p_R^t)/[(r_R^t - r_A^t) p_R^t]$. If $\delta - r_A^t p_A^t - r_H p_H^t - r_A^t p_R^t < 0$, we have $q_t^\star(x_t, p_t) = 0$.

If the right-hand side of this inequality is within $[0, 1]$, i.e., $p_H^t r_H + (1 - p_H^t) r_A^t \leq \delta \leq p_R^t r_R^t + p_H^t r_H + p_A^t r_A^t$, we have

$$q_t^\star(x_t, p_t) = \frac{\delta - r_A^t p_A^t - r_H p_H^t - r_A^t p_R^t}{(r_R^t - r_A^t) p_R^t} = \frac{\delta - r_A^t (p_A^t + p_R^t) - r_H p_H^t}{(r_R^t - r_A^t) p_R^t} = \frac{\delta - r_A^t + (r_A^t - r_H) p_H^t}{(r_R^t - r_A^t) p_R^t}.$$

If the right-hand side is larger than 1, i.e., $\delta > p_R^t r_R^t + p_H^t r_H + p_A^t r_A^t$, the problem is infeasible.

Summing up the analysis above, we have

$$q_t^\star(x_t, p_t) = \begin{cases} \dfrac{\delta - r_A^t + (r_A^t - r_H) p_H^t}{(r_R^t - r_A^t) p_R^t}, & \text{if } p_R^t \neq 0,\ \delta \geq p_H^t r_H + (1 - p_H^t) r_A^t \\ & \text{and } \delta \leq p_R^t r_R^t + p_H^t r_H + p_A^t r_A^t, \\ 0, & \text{if } p_R^t \neq 0 \text{ and } \delta < p_H^t r_H + (1 - p_H^t) r_A^t, \\ & \text{or } p_R^t = 0 \text{ and } \delta \leq p_H^t r_H + (1 - p_H^t) r_A^t \\ \text{infeasible}, & \text{if } p_R^t \neq 0 \text{ and } \delta > p_R^t r_R^t + p_H^t r_H + p_A^t r_A^t, \\ & \text{or } p_R^t = 0 \text{ and } \delta > p_H^t r_H + (1 - p_H^t) r_A^t. \end{cases}$$

Equivalently, on the feasible domain,

$$p_R^t q_t^\star(x_t, p_t) = \left[\frac{\delta - r_A^t + (r_A^t - r_H) p_H^t}{r_R^t - r_A^t}\right]_+. \tag{18}$$

For a prescribed feasible path $p_H(s)$ in a region with $r_A(x) > r_H$ and $r_R(x) > r_A(x)$, a formal continuous-time analogue of (3) is

$$\frac{dx(s)}{ds} = \left[\frac{\delta - r_A(x(s)) + (r_A(x(s)) - r_H) p_H(s)}{r_R(x(s)) - r_A(x(s))}\right]_+.$$

This replaces the per-stage learning increment by a rate; justifying a limiting approximation requires an explicit time-step scaling and additional regularity. We retain the discrete-time recursion to analyze stagewise capacity and automation thresholds.

F.2 PROOF FOR LEMMA 1: MONOTONICITY OF VALUE FUNCTION

We prove this lemma by induction. We first record a useful implication of Assumption 2. Define

$$z(x, p_H) = \left[\frac{\delta - r_A(x) + (r_A(x) - r_H) p_H}{r_R(x) - r_A(x)}\right]_+.$$

On the feasible domain, $z(x, p_H)$ is nonincreasing in x. To see this, consider the interior region in which the term inside $[\cdot]_+$ is positive. Let $a = r_A(x)$ and $r = r_R(x)$. Then $z = [\delta - a + (a - r_H)p_H]/(r - a)$. For a feasible policy, $z \le 1 - p_H$, which is equivalent to $\delta \le p_H r_H + (1 - p_H)r$. Thus,

$$\frac{\partial z}{\partial a} = \frac{\delta - \{p_H r_H + (1 - p_H)r\}}{(r - a)^2} \le 0, \qquad \frac{\partial z}{\partial r} = -\frac{\delta - a + (a - r_H)p_H}{(r - a)^2} \le 0.$$

Since both $r_A(x)$ and $r_R(x)$ are nondecreasing in x, $z(x, p_H)$ is nonincreasing in x on the feasible domain. If the term inside $[\cdot]_+$ is nonpositive, then $z(x, p_H) = 0$ and the same conclusion holds.

By definition, $V_{T+1}(x_{T+1}) \equiv 0$ is nondecreasing with respect to x_{T+1}. Suppose $V_{t+1}(x_{t+1})$ is nondecreasing with respect to x_{t+1}. We show $V_t(x_t)$ is nondecreasing with respect to x_t. Take any two states $x' \ge x$ and any policy $p_t \in \mathcal{P}(x) = \{p_t \in \Delta^3 : z(x, p_H^t) \in [0, p_R^t], c_H p_H^t + c_R z(x, p_H^t) \le b\}$. From the argument above, $z(x', p_H^t) \le z(x, p_H^t)$. Therefore the human effort constraint and the feasibility condition remain satisfied at x', so $p_t \in \mathcal{P}(x')$.

From (18), the reward after substituting the human best response can be written as $\tilde{\ell}(x, p_t) = \max\{\delta,\ p_H^t r_H + (1 - p_H^t) r_A(x)\}$ on the feasible domain. Since $r_A(x)$ is nondecreasing, $\tilde{\ell}(x, p_t)$ is nondecreasing in x. Define $p^\star(x) \triangleq \arg\max_{p \in \mathcal{P}(x)} \tilde{\ell}(x, p)$. From the Bellman equation (7),

$$V_t(x) \;=\; \max_{p_t \in \mathcal{P}(x)} \Big\{ \tilde{\ell}(x, p_t) + V_{t+1}\big(x + \tilde{g}(x, p_t)\big) \Big\} \;=\; \tilde{\ell}(x, p_t^\star) + V_{t+1}\big(x + \tilde{g}(x, p_t^\star)\big)$$

Moreover, by Assumption 3, the post-decision learning state $x + \tilde{g}\big(x, p_t^\star(x)\big) = x + \lambda p_{t,H}^\star(x) \mathbf{1}\{r_A(x) \le r_H\} + z\big(x, p_{t,H}^\star(x)\big)$ is nondecreasing in x. Combining this with the induction hypothesis, we have $V_{t+1}\big(x' + \tilde{g}(x', p_t)\big) \ge V_{t+1}\big(x + \tilde{g}(x, p_t)\big)$. Thus, every policy feasible at x yields weakly higher Bellman payoff at x'. Since $\mathcal{P}(x) \subseteq \mathcal{P}(x')$, taking the maximum over feasible policies gives $V_t(x') \ge V_t(x)$. The result follows by induction.

F.3 Proof for Lemma 2: Feasible Region

Fix a learning state x_t. Define the state-dependent feasible set

$$\mathcal{I}(x_t) \triangleq \Big\{ (\delta, b) : \exists (p_t, q_t) \in \Delta^3 \times [0, 1] \text{ s.t. } c_H p_H^t + c_R p_R^t q_t \le b \text{ and } \ell(x_t, p_t, q_t) \ge \delta \Big\},$$

where ℓ is given in (4). This is the same notion as the single-stage feasible region in Section A.1, with the fixed benchmarks (r_A, r_R) replaced by $(r_A(x_t), r_R(x_t))$.

For fixed x_t, let $h = p_H^t$ and $s = p_R^t q_t$ denote the fraction of human-only decisions and the effective reviewed fraction under assisted decision making. Since $s \le p_R^t$ and $p_A^t + p_H^t + p_R^t = 1$, the feasible set can be written as $h \ge 0$, $s \ge 0$, $h + s \le 1$, $c_H h + c_R s \le b$. For any such pair (h, s), the induced reward is $r_A(x_t)(1 - h - s) + r_H h + r_R(x_t) s$. Therefore the maximal feasible single-stage reward at state x_t is

$$M(x_t; b) = \max_{\substack{h, s \ge 0 \\ h + s \le 1 \\ c_H h + c_R s \le b}} \{ r_A(x_t)(1 - h - s) + r_H h + r_R(x_t) s \}.$$

Consequently, $\mathcal{I}(x_t) = \{(\delta, b) : \delta \le M(x_t; b)\}$.

Under Assumption 2, $r_A(x)$ and $r_R(x)$ are nondecreasing in x. Hence, for any fixed (h, s), the reward $r_A(x)(1 - h - s) + r_H h + r_R(x) s$ is nondecreasing in x. Taking the maximum over the same feasible set implies that $M(x; b)$ is nondecreasing in x. Since the learning state satisfies $x_t \ge x_0$ for every t, we have $M(x_t; b) \ge M(x_0; b)$ for all $t \in \{0, 1, \ldots, T\}$. Therefore, the most restrictive feasibility condition is the initial one. The multi-stage problem is feasible if the stage problem is feasible at any stage. Since $\mathcal{I}(x_0) \subseteq \mathcal{I}(x_t)$ for every t, the horizon-level feasible region is $\mathcal{I}(\{x_t\}_{t=0}^T) = \bigcap_{t=0}^T \mathcal{I}(x_t) = \mathcal{I}(x_0) = \{(\delta, b) : \delta \le M(x_0; b)\}$.

According to Section A.1, for a fixed x_0, the following boundaries characterize the feasible region $\{(\delta, b) : \delta \le M(x_0; b)\}$ more specifically:

1. Performance upper boundary: Because $\ell(x_0, p_0, q_0) \le r_R^0$ for all (p_0, q_0), every point with $\delta > r_R^0$ is infeasible. Thus the upper boundary is the horizontal line $\delta = r_R^0$.
2. Universal-assistance (sufficient) boundary: For $\delta \in (r_A^0, r_R^0]$, consider universal assistance $p_R^0 = 1$. The minimal modification intensity needed to meet δ is $q_{\min}(x_0, \delta) = (\delta - r_A^0)/(r_R^0 - r_A^0)$, so universal assistance is feasible if

$$b \ge c_R^0\, q_{\min}^0(x_0, \delta) = \frac{c_R}{r_R^0 - r_A^0}\,(\delta - r_A^0) \Leftrightarrow \delta \le r_A^0 + \frac{r_R^0 - r_A^0}{c_R} b.$$

Summarizing (i) and (ii), for a fixed initial learning state x_0, the feasible set is defined as follows

$$\mathcal{I}(x_0) \triangleq \Big\{(b, \delta) : \delta \le r_R^0 \text{ and } \delta \le r_A^0 + \frac{r_R^0 - r_A^0}{c_R} b\Big\}.$$

F.4 Proof for Proposition 8: Existence of the Multi-stage Equilibrium

Since the feasibility conditions in Lemma 2 are satisfied, the constraint set of the multi-stage model (6) is nonempty. Recall that the objective function of (6) is to maximize $L(\{p_t\}_{t=0}^T) = \sum_{t=0}^T \ell\big(x_t, p_t, q_t^\star(x_t, p_t)\big)$.

In order to show that the single-agent control problem (6) admits a solution, we use Weierstrass' Theorem (e.g., Proposition 2.1.1 of Bertsekas et al. (2003)). Consider a closed proper function $-L : \{P(x_t)\}_{t=0}^T \to (-\infty, \infty]$. If its effective domain, $\mathrm{dom}(-L) \triangleq \Big\{\big(p_t \in \mathcal{P}(x_t)\big)_{t=0}^T : -L(\{p_t\}_{t=0}^T) < \infty\Big\}$, is bounded, then the set of minima of $-L$ over the feasible set $\{P(x_t)\}_{t=0}^T$ is nonempty and compact.

We need to verify the conditions of Weierstrass' Theorem. First, $\mathrm{dom}(-L)$ is bounded because $p_t \in \Delta^3$ for each t. Moreover, the learning increment is bounded by $\lambda + 1$, and hence $x_t \le t(\lambda + 1)$. Since $r_A(\cdot)$ and $r_R(\cdot)$ are continuous, they are bounded on the compact set of reachable states. Thus $-L$ is proper.

It remains to show that $-L$ is a closed function. According to Proposition 1.2.3 of Bertsekas et al. (2003), it suffices to show that (i) L is upper semicontinuous over $\mathrm{dom}(L)$ and (ii) $\mathrm{dom}(-L)$ is closed.

Step 1: We show that the post-decision state is upper semicontinuous. From (18), on the feasible domain,

$$p_R^t q_t^\star(x_t, p_t) = z(x_t, p_H^t) = \left[\frac{\delta - r_A^t + (r_A^t - r_H)p_H^t}{r_R^t - r_A^t}\right]_+.$$

Since $r_A(\cdot)$ and $r_R(\cdot)$ are continuous, the mapping $(x_t, p_t) \mapsto z(x_t, p_H^t)$ is continuous. The post-decision state can be written as $\Phi(x_t, p_t) = x_t + \tilde{g}(x_t, p_t) = x_t + z(x_t, p_H^t) + \lambda p_H^t \mathbf{1}\{r_A(x_t) \le r_H\}$. The set $\{x_t \ge 0 : r_A(x_t) \le r_H\}$ is closed because $r_A(\cdot)$ is continuous, and the indicator of a closed set is upper semicontinuous. Therefore, $(x_t, p_t) \mapsto \Phi(x_t, p_t)$ is upper semicontinuous. By Assumption 3, it is also nondecreasing in x_t.

Step 2: Denote $\mathbf{p} \triangleq (p_0, p_1, \cdots, p_T)$. We show that for each $t \in \{0, 1, \ldots, T\}$, $\mathbf{p} \mapsto x_t(\mathbf{p})$ is upper semicontinuous. The claim holds for $t = 0$ since $x_0 \equiv 0$. Suppose it holds for t. Since $x_{t+1}(\mathbf{p}) = \Phi(x_t(\mathbf{p}), p_t)$, and since Φ is upper semicontinuous and nondecreasing in its state argument, the composition $\mathbf{p} \mapsto x_{t+1}(\mathbf{p})$ is upper semicontinuous (e.g., Exercise 1.40 of Rockafellar & Wets (2009)). Thus the claim follows by induction.

Step 3: We show $\mathbf{p} \mapsto \tilde{\ell}(x_t(\mathbf{p}), p_t)$ is upper semicontinuous for each $t \in \{0, 1, \ldots, T\}$. Consequently, $\mathbf{p} \mapsto L(\mathbf{p})$ is also upper semicontinuous. Let $\tilde{\ell}(x_t, p_t) \triangleq \ell(x_t, p_t, q_t^\star(x_t, p_t))$. From (18), we have

$$\tilde{\ell}(x_t, p_t) = \begin{cases} \delta, & \text{if } \delta \ge p_H^t r_H + (1 - p_H^t) r_A^t, \\ p_H^t r_H + (1 - p_H^t) r_A^t, & \text{otherwise.} \end{cases}$$

Hence $\tilde{\ell}(x_t, p_t)$ is continuous and nondecreasing in x_t. Since $\mathbf{p} \mapsto x_t(\mathbf{p})$ is upper semicontinuous, the composition $\mathbf{p} \mapsto \tilde{\ell}(x_t(\mathbf{p}), p_t)$ is upper semicontinuous (e.g., Exercise 1.40 of Rockafellar & Wets (2009)). Therefore, $L(\mathbf{p}) = \sum_{t=0}^T \tilde{\ell}(x_t(\mathbf{p}), p_t)$ is upper semicontinuous as a finite sum of upper semicontinuous functions.

Step 4: It remains to show that the effective domain of $-L$, $\mathrm{dom}(-L)$, is closed. For state x_t and action $p_t = (p_A^t, p_H^t, p_R^t) \in \Delta^3$, the feasibility constraints can be written as

$$z(x_t, p_H^t) \le p_R^t, \qquad c_H p_H^t + c_R z(x_t, p_H^t) \le b.$$

As shown in the proof of Lemma 1, $z(x_t, p_H^t)$ is continuous and nonincreasing in x_t on the feasible domain. Take any sequence $\mathbf{p}^{(n)} \to \mathbf{p}$ with each $\mathbf{p}^{(n)}$ feasible. From Step 2, $x_t(\mathbf{p})$ is upper semicontinuous in $\mathbf{p}$, so $x_t(\mathbf{p}) \ge \limsup_{n\to\infty} x_t(\mathbf{p}^{(n)})$. Since z is nonincreasing in the state and continuous in p_H^t, this implies the composition $\mathbf{p} \to z(x_t(\mathbf{p}), p_H^t)$ is lower semicontinuous (l.s.c.) (e.g., Exercise 1.40 of Rockafellar & Wets (2009)). Thus, by definition of l.s.c., $z(x_t(\mathbf{p}), p_H^t) \le \liminf_{n\to\infty} z(x_t(\mathbf{p}^{(n)}), p_H^{t,(n)})$. Because each $\mathbf{p}^{(n)}$ is feasible,

$$z(x_t(\mathbf{p}^{(n)}), p_H^{t,(n)}) \le p_R^{t,(n)}, \qquad c_H p_H^{t,(n)} + c_R z(x_t(\mathbf{p}^{(n)}), p_H^{t,(n)}) \le b.$$

Taking limits gives $z(x_t(\mathbf{p}), p_H^t) \le p_R^t$, $c_H p_H^t + c_R z(x_t(\mathbf{p}), p_H^t) \le b$. Therefore, the feasibility constraints are preserved under limits, and the feasible policy sequence set is closed. Consequently, $\mathrm{dom}(-L)$ is closed.

The conditions of Weierstrass' Theorem are satisfied. Hence the multi-stage model (6) admits an optimal policy sequence, which gives an equilibrium of the coordinator–human interaction after substituting the human's stage-wise best response.

F.5 Propositions for Region with Dynamic Policies

In this section, we summarize the formal propositions for the dynamic equilibrium regions in Table 2. We first consider the case in which human-only decisions generate at least as much learning signal as assisted review, namely $\lambda \ge 1$. In this case, the static and dynamic incentives are aligned in the two low-requirement regions: the same human allocation that maximizes current payoff also generates the strongest learning signal.

Proposition 9 (Human-only phase followed by automation)**.** *Suppose Assumptions 2 and 3 hold, $r_A^0 < r_H$, $\delta \le r_H$, $b \ge c_H$, and $\lambda \ge 1$. Then an equilibrium policy is*

$$(p_A^t, p_H^t, p_R^t, q_t) = \begin{cases} (0,1,0,0), & r_A^t < r_H, \\ (1,0,0,0), & r_A^t \ge r_H. \end{cases}$$

Proposition 10 (Mixed human and AI phase followed by automation)**.** *Suppose Assumptions 2 and 3 hold, $r_A^0 < r_H$, $b < c_H$, $\delta \le r_A^0 + \frac{b}{c_H}(r_H - r_A^0)$, and $\lambda \ge 1$. Then an equilibrium policy is*

$$(p_A^t, p_H^t, p_R^t, q_t) = \begin{cases} \left(1 - \frac{b}{c_H}, \frac{b}{c_H}, 0, 0\right), & r_A^t < r_H, \\ (1,0,0,0), & r_A^t \ge r_H. \end{cases}$$

Propositions 9 and 10 show that, when $\lambda \ge 1$, the coordinator uses the deployable human capacity exactly as in the corresponding static region until the AI reaches r_H. Greater deployable human capacity increases the amount of human-only processing before that crossing and therefore accelerates the transition to automation.

When $\lambda < 1$, assisted review creates more learning per unit of engaged mass than human-only processing. The coordinator may then use the Bellman solution in early states, because the policy that is best for current reward need not be best for continuation value. Define $\kappa_t = (r_H - \lambda r_R^t)/(1-\lambda)$. Above this threshold, the dynamic value of assisted review is no longer sufficient to displace the static allocation. Propositions 9 and 11 are proved in Appendix F.6.2.

Proposition 11 (Bellman phase followed by human-only processing and automation)**.** *Suppose Assumptions 2 and 3 hold, $r_A^0 < r_H$, $\delta \le r_H$, $b \ge c_H$, and $\lambda < 1$. Then an equilibrium policy satisfies*

$$(p_A^t, p_H^t, p_R^t, q_t) = \begin{cases} \textit{a solution of } (7), & r_A^t < \kappa_t, \\ (0,1,0,0), & \kappa_t \le r_A^t < r_H, \\ (1,0,0,0), & r_A^t \ge r_H. \end{cases}$$

Proposition 12 (Bellman phase followed by mixed processing and automation)**.** *Suppose Assumptions 2 and 3 hold, $r_A^0 < r_H$, $b < c_H$, $\delta \le r_A^0 + \frac{b}{c_H}(r_H - r_A^0)$, and $\lambda < 1$. Then an equilibrium policy satisfies*

$$(p_A^t, p_H^t, p_R^t, q_t) = \begin{cases} \textit{a solution of } (7), & r_A^t < \kappa_t, \\ \left(1 - \frac{b}{c_H}, \frac{b}{c_H}, 0, 0\right), & \kappa_t \le r_A^t < r_H, \\ (1,0,0,0), & r_A^t \ge r_H. \end{cases}$$

The remaining regions are governed by the performance requirement. Consider first the case in which the requirement is above what the initial mixed allocation can satisfy, but no higher than the human benchmark. For $b < c_H$, define $\theta(\delta, b) = (c_H\delta - br_H)/(c_H - b)$, and define $\rho_t = \theta(\delta, b)$ if $\lambda \ge 1$, while $\rho_t = \max\{\theta(\delta, b), \kappa_t\}$ if $\lambda < 1$.

Proposition 13 (Bellman phase followed by mixed processing and automation)**.** *Suppose Assumptions 2 and 3 hold, $\delta \le M(x_0; b)$, $r_A^0 < r_H$, and $r_A^0 + b/c_H(r_H - r_A^0) < \delta \le r_H$. Then $b < c_H$, and an equilibrium policy satisfies*

$$(p_A^t, p_H^t, p_R^t, q_t) = \begin{cases} \textit{a solution of } (7), & r_A^t < \rho_t, \\ \left(1 - \frac{b}{c_H}, \frac{b}{c_H}, 0, 0\right), & \rho_t \le r_A^t < r_H, \\ (1,0,0,0), & r_A^t \ge r_H. \end{cases}$$

The threshold $\theta(\delta, b)$ is the AI performance level at which the capacity-constrained mixed allocation first satisfies the requirement. It is increasing in δ and decreasing in b. Stricter standards therefore delay the mixed phase, while greater deployable human capacity make it attainable earlier.

Finally, suppose the requirement exceeds the human benchmark. Human-only processing cannot be a terminal operating mode in this region, because it does not meet the requirement. The terminal threshold is therefore δ, not r_H.

Proposition 14 (Bellman phase followed by automation under a high requirement)**.** *Suppose Assumptions 2 and 3 hold, $\delta \le M(x_0; b)$, $\delta \ge r_H$, and $r_A^0 \le \delta$. Then an equilibrium policy satisfies*

$$(p_A^t, p_H^t, p_R^t, q_t) = \begin{cases} \textit{a solution of } (7), & r_A^t < \delta, \\ (1,0,0,0), & r_A^t \ge \delta. \end{cases}$$

Proposition 14 shows that when the performance requirement exceeds the human benchmark, even fully relying on humans is insufficient for long-run feasibility. Automation becomes viable only once AI performance itself crosses the required threshold. Until that point, human involvement serves primarily as a learning mechanism rather than a permanent solution. The timing of automation is therefore governed directly by the learning curve and the required performance level: stricter requirements push the transition further into the future.

F.6 Proofs for Region with Dynamic Policies

In this section, we provide the proofs of the propositions analyzing the dynamic policies in Appendix F.5. We begin by introducing Lemma 9.

Lemma 9. *Suppose that, at some stage $\hat{t}$, we have $r_A^{\hat{t}} \ge \max\{r_H, \delta\}$. Then for stages $t \ge \hat{t}$, the equilibrium policy is $p_A^t = 1$, $p_H^t = 0$, $p_R^t = 0$, $q_t = 0$.*

A direct implication of Lemma 9 is the following: if the feasibility conditions in Lemma 2 are satisfied, $r_A^0 \ge r_H$, and $\delta \le r_H$, then the equilibrium policy for the multi-stage model (6) is full automation, i.e., $p_A^t = 1, p_H^t = 0, p_R^t = 0, q_t = 0$ for all $t \in \{0, 1, 2, \cdots, T\}$. This lemma is proved in Appendix F.6.1.

F.6.1 Proof of Lemma 9

Since $r_A^t \geq r_H$ for all stages $t \geq \hat{t}$, from the coordinator's perspective, we have $p_H^t = 0$ for stages $t \geq \hat{t}$. Thus, the human's optimal policy at stage $t \geq \hat{t}$ in Appendix F.1 reduces to

$$q_t^\star(x_t, p_t) = \begin{cases} \dfrac{\delta - r_A^t}{(r_R^t - r_A^t)p_R^t}, & \text{if } p_R^t \neq 0,\ r_A^t \leq \delta \leq p_R^t r_R^t + (1 - p_R^t) r_A^t, \\ 0, & \text{if } p_R^t \neq 0 \text{ and } \delta < r_A^t, \text{or } p_R^t = 0 \text{ and } \delta \leq r_A^t. \end{cases}$$

Since $r_A^t \geq \delta$ for all stages $t \geq \hat{t}$, we must have $q_t^\star(x_t, p_t) = 0$ for stages $t \geq \hat{t}$.

Given $q_t^\star(x_t, p_t) = 0$ for stages $t \geq \hat{t}$, we next show that $p_H^t = 0$ for stages $t \geq \hat{t}$ is indeed optimal for the coordinator. From the coordinator's perspective, the expected stage-t reward (4) reduces to $\hat{\ell}(x_t, p_t, q_t^\star(x_t, p_t)) = r_A^t p_A^t + r_A^t p_R^t + r_H p_H^t$ for all $t \in \{\hat{t}, \hat{t}+1, \ldots, T\}$, and the per-stage deployable human-capacity constraint becomes $c_H p_H^t \leq b$ for all $t \in \{\hat{t}, \hat{t}+1, \ldots, T\}$. For $t \geq \hat{t}$, options R and A are equivalent because they yield the same reward r_A^t and use the same amount of per-stage human capacity. Moreover, option H is weakly dominated by options R and A because $r_A^t \geq r_H$ and option H incurs cost c_H. Thus, $p_H^t = p_R^t = 0$ and $p_A^t = 1$ constitute the equilibrium policy for the multi-stage model (6) at stage $t \geq \hat{t}$. Note that Proposition 14 is a direct application of Lemma 9.

F.6.2 Proof of Proposition 9 and Proposition 11: (Dynamic Coordination →) Human-Only → AI-Only

From the Bellman recursion (7) of the multi-stage model (6), at each stage $t \in \{0, 1, \ldots, T\}$,

$$V_t(x_t) = \max_{p_t \in \mathcal{P}(x_t)} \Big\{ \tilde{\ell}(x_t, p_t) + V_{t+1}\big(x_t + \tilde{g}(x_t, p_t)\big) \Big\}. \tag{19}$$

Based on the human's closed-form solution, $q_t^\star(x_t, p_t)$, as in Appendix F.1, the function $\tilde{g}(x_t, p_t)$ in the state transition (3) can be computed as

$$\tilde{g}(x_t, p_t) = g\big(x_t, p_t, q_t^\star(x_t, p_t)\big) = \begin{cases} \lambda p_H^t + \left[\frac{\delta - r_A^t + (r_A^t - r_H)p_H^t}{r_R^t - r_A^t}\right]_+, & \text{if } r_A^t \leq r_H, \\ \left[\frac{\delta - r_A^t + (r_A^t - r_H)p_H^t}{r_R^t - r_A^t}\right]_+, & \text{if } r_A^t > r_H. \end{cases}$$

where $[z]_+ \triangleq \max\{z, 0\}$. First, we derive the conditions under which $\tilde{g}(x_t, p_t)$ will not decrease with respect to p_H^t.

(i) When $r_A^t > r_H$, $\tilde{g}(x_t, p_t) = \left[\frac{\delta - r_A^t + (r_A^t - r_H)p_H^t}{r_R^t - r_A^t}\right]_+$, which is nondecreasing in p_H^t since $r_R^t > r_A^t$.

(ii) When $r_A^t \leq r_H$ and $\delta - r_A^t + (r_A^t - r_H)p_H^t \leq 0$, $\tilde{g}(x_t, p_t) = \lambda p_H^t$, which is strictly increasing in p_H^t since $\lambda > 0$.

(iii) When $r_A^t \leq r_H$ and $\delta - r_A^t + (r_A^t - r_H)p_H^t > 0$, we have

$$\tilde{g}(x_t, p_t) = \lambda p_H^t + \frac{\delta - r_A^t + (r_A^t - r_H)p_H^t}{r_R^t - r_A^t} = \frac{[\lambda(r_R^t - r_A^t) + (r_A^t - r_H)]p_H^t + \delta - r_A^t}{r_R^t - r_A^t}.$$

Since $r_R^t > r_A^t$, the function $\tilde{g}(x_t, p_t)$ is nondecreasing with respect to p_H^t when $\lambda \geq (r_H - r_A^t)/(r_R^t - r_A^t)$.

From the analysis above, when $\lambda \geq (r_H - r_A^t)/(r_R^t - r_A^t)$, $\tilde{g}(x_t, p_t)$ is nondecreasing in p_H^t. By Lemma 1, $V_{t+1}\big(x_t + \tilde{g}(x_t, p_t)\big)$ is also nondecreasing in p_H^t for fixed x_t.

(1) When $r_A^t < r_H$, $\delta \leq r_H$, and $b \geq c_H$: If $\lambda \geq (r_H - r_A^t)/(r_R^t - r_A^t)$, then for the first part of (19), from Theorem 1 and Case 1.4 in Appendix A.2, the equilibrium policy for $\max_{p_t \in \mathcal{P}(x_t)} \tilde{\ell}(x_t, p_t)$ is $p_H^t = 1$ and $p_R^t = p_A^t = q_t = 0$. Since $V_{t+1}\big(x_t + \tilde{g}(x_t, p_t)\big)$ is nondecreasing in p_H^t for fixed x_t, this equilibrium policy also maximizes the second part of (19). As a result, this equilibrium policy maximizes V_t and is the equilibrium policy for the multi-stage model (6) at stage t.

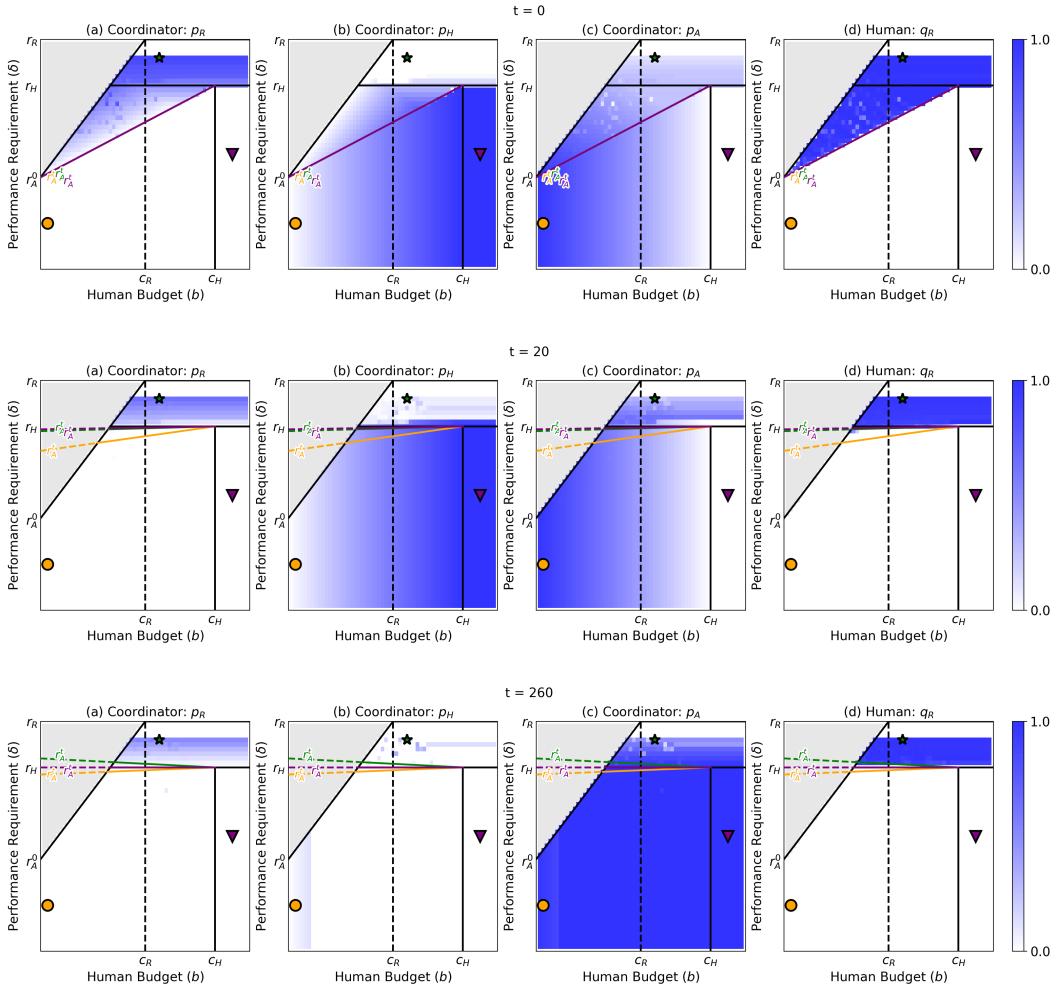


Figure 8: Evolution of the human and coordinator policies in the multi-stage model when $t = 0, 20, 260$. The intensity of blue represents the corresponding policy probability. We illustrate the scenarios corresponding to three representative (δ, b) points, along with their associated r_A^t over time: the orange circle in the *no-collaboration* regime, the purple triangle in the *human-only* regime, the green star in the *full collaboration* regime.

It remains to figure out when $\lambda \geq (r_H - r_A^t)/(r_R^t - r_A^t)$ holds. When $\lambda \geq 1$, this inequality holds since $r_H \leq r_R^t$. When $\lambda < 1$, the condition is equivalent to $r_A^t \geq (r_H - \lambda r_R^t)/(1-\lambda)$. If $\lambda < (r_H - r_A^t)/(r_R^t - r_A^t)$, i.e., $\lambda < 1$ and $r_A^t < (r_H - \lambda r_R^t)/(1-\lambda)$, there is no closed-form solution for the equilibrium policy, and we need to solve the model using dynamic programming.

(2) When $r_A^t \geq r_H$, $\delta \leq r_H$, and $b \geq c_H$, from Lemma 9, the equilibrium policy for the multi-stage model (6) is $p_A^t = 1$ and $p_H^t = p_R^t = q_t = 0$.

The proofs for Propositions 10, 12, and 13 are similar and thus omitted.

F.7 Evolution of the Dynamic Equilibrium Policies

Relative to the core (single-stage) model, the key new feature is that coordination decisions influence the evolution of AI performance, so that feasibility and equilibrium policies must be understood jointly with the learning dynamics. Nevertheless, as we show below using the numerical experiments described in Appendix G, the geometric structure of the policy regions remains closely related to the static case: at any given time, the (δ, b) space is partitioned by the same set of boundaries, but these boundaries shift endogenously as the AI improves.

Fig. 8 illustrates how the equilibrium policies evolve across representative time stages. Each subplot in Fig. 8 corresponds to Fig. 5a or Fig. 6 of the core (single-stage) model, which demonstrates the equilibrium policy for both the coordinator and the human at a given time stage $t \in \{1, 2, \cdots, T\}$. Across all panels, the qualitative regime structure is preserved. What changes over time is the effective AI baseline r_A^t, which increases as the AI learns from human involvement. In Fig. 8, we show how r_A^t evolves for scenarios corresponding to three representative (δ, b) points: the orange circle in the *no-collaboration* regime, the purple triangle in the *human-only* regime, the green star in the *full collaboration* regime. As r_A^t rises, the region in which AI alone can meet the performance requirement expands, and the portion of the space requiring human-only involvement correspondingly contracts. Thus, the dynamic model can be interpreted as a sequence of static problems with an improving baseline, linked together by the learning process induced by human participation.

These patterns reveal two central mechanisms. First, human involvement generates an intertemporal learning externality: allocating tasks to humans early on improves the AI, which reduces the need for human intervention later. Second, the performance requirement and the deployable human capacity jointly determine not only whether oversight is used at a given time but also the speed of the transition toward automation. Stricter performance requirements sustain human involvement for longer horizons, whereas greater deployable human capacity accelerate learning and lead to earlier convergence to AI-dominant policies. As we can observe from Fig. 8, the green r_A^t increases most rapidly due to the highest performance requirement, whereas the growth of the orange r_A^t is the slowest since the deployable human capacity is limited and the performance requirement is low.

G Configuration for Synthetic Experiments

Core Model The parameters of the core model (1) (2) are listed as follows: $r_H = 0.80, r_R = 1.00$, $r_A = 0.40$; $c_H = 1.00$, $c_R = 0.60$; $\delta \in (0, 1.00)$, $b \in (0, 1.20]$. Fig. 5a and Fig. 5b are generated using these parameters.

Multi-stage Model Motivated by statistical learning theory and empirical scaling laws, we assume AI error decays with accumulated supervision and capture this with a simple power-law specification for the AI baseline:

$$r_A^t \triangleq r_A(x_t) = r_R - \frac{\alpha}{1 + x_t^{\beta}}, \qquad \text{given } \alpha > 0, \beta > 0, \tag{20}$$

Under (20), r_A^t is increasing in x_t and converges to r_R as $x_t \to \infty$; in particular $r_A^0 = r_R - \alpha$ and α measures the initial gap to the assisted outcome. The parameter β controls the concavity (learning speed). The functional form is a convenient, empirically supported shorthand for the common observation that predictive error falls with additional labeled data, often approximately following a power law (Vapnik, 1998; Shalev-Shwartz & Ben-David, 2014; Hestness et al., 2017; Kaplan et al., 2020). For transparency, we assume r_R is the asymptotic upper bound on performance attainable with sufficient supervision. This normalization simplifies exposition; none of the qualitative results or the equilibrium phenomena rely on r_R being the exact attainable ceiling.

The model parameters of the multi-stage model (6) are summarized as follows: $T = 300$, $\alpha = 0.60$, $\beta = 0.20$, $r_H = 0.80$, $r_R = 1.00$, $c_H = 1.00$, $c_R = 0.60$; $\delta \in (0, 1.00)$, $b \in (0, 1.20]$; $\lambda = 1$. Fig. 5b and Fig. 8 are derived under this parameter setting.

We use the following numerical settings to implement the Bellman recursion (7): (i) State discretization: We discretize the state x using a uniform grid on $[0, T]$. The value function $V_t(x)$ is represented at each grid point. In the experiments, we used 801 grid points in total for $T = 300$. (ii) Value function approximation: For off-grid state x', the continuation value is approximated by linear interpolation using the nearest grid states x_k and x_{k+1}: $V_{t+1}(x') \approx (1-w)\, V_{t+1}(x_k) + w\, V_{t+1}(x_{k+1})$, where $w = \frac{x'-x_k}{x_{k+1}-x_k}$, $x_k \le x' \le x_{k+1}$. (iii) Action discretization: Similar to the state discretization, we discretize the continuous action space (p_H^t, p_R^t, p_A^t). In particular, we enumerate a mesh of (p_H, p_R) pairs with a step size (default 0.01), and set $p_A = 1 - p_H - p_R$. For computational tractability, this approximation replaces the nonlinear program at each iteration with a grid search over the feasible region. (iv) Numerical tolerance: The numerical tolerance is set to 10^{-12} throughout.

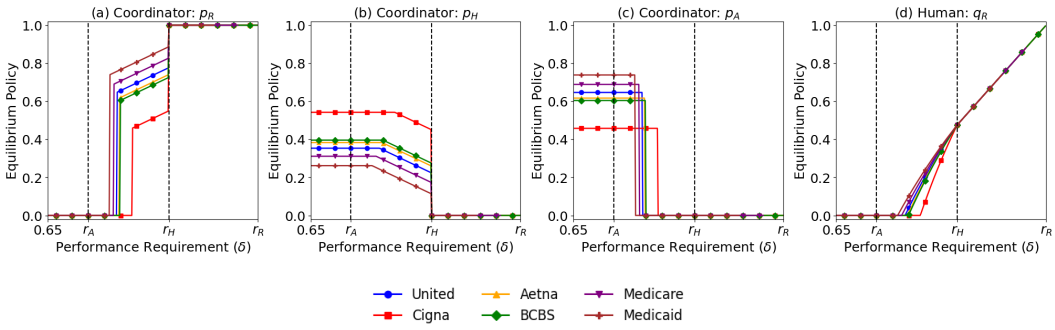


Figure 9: Calibrated model illustration of the equilibrium policy under different performance requirements.

Heterogeneous Multi-task Model The model inputs of the multi-task model (16) are summarized as below: $\lambda_1 = \lambda_2 = 0.5$, $v_1 = 1.00$, $v_2 = 1.25$; $r_{1H} = 0.80$, $r_{2H} = 0.88$, $r_{1R} = 0.91$, $r_{2R} = 0.96$, $r_{1A} = 0.70$, $r_{2A} = 0.76$; $c_{1H} = \$105$, $c_{2H} = \$150$, $c_{1R} = \$37$, $c_{2R} = \$60$; $\delta_1 \in (0.65, 0.95)$, $\delta_2 \in (0.70, 1.00)$; $b \in \{\$30, \$45, \$65, \$95\}$. Fig. 7 is derived using these parameters.

H Configuration for Empirical Study of Radiographic Diagnosis and Additional Results

The purpose of the calibration is not to estimate a structural model of radiology operations, but to map empirically plausible performance and cost levels into the equilibrium regions characterized. We use empirically informed benchmark scenarios for r_A, r_H, and r_R, and we use reimbursement as a proxy for deployable human capacity in dollar-equivalent resource units per case. Under this interpretation, variation across payers captures differences in the provider's ability to sustain radiologist time after an AI recommendation has been generated.

For diagnostic performance, we take human radiologist performance to be approximately 80% expected accuracy (Patel et al., 2019). Standalone AI performance varies across systems. CheXNet achieves approximately 70% accuracy, whereas the more recent CheXMax system achieves approximately 84% accuracy (Patel et al., 2019). The actively reviewed human–AI condition has the highest reported performance, approximately 91%, which we use as the active-collaboration benchmark r_R; the model's delivered assisted-mode performance still depends on q_R. On the cost side, a fully human radiology workflow costs approximately \$105 per case, while an AI-assisted workflow costs approximately \$37 per case (El Arab & Al Moosa, 2025; MDsave, 2026). We compare these costs with payer-specific reimbursement levels. Commercial insurers such as UnitedHealthcare, Cigna, Aetna, and Blue Cross Blue Shield reimburse approximately \$37 to \$57 per case (Payerprice, 2026), whereas public programs reimburse less, with Medicare at approximately \$33 and Medicaid at approximately \$27 per case (MediBillMD, 2025; HCS, 2026).

Fig. 9 shows the same mechanism like Fig. 4 from the perspective of the performance requirement. When δ is below the standalone AI benchmark, AI-only diagnosis can satisfy the requirement, so assisted review is unnecessary. Once δ exceeds r_A, payers with lower calibrated deployable human capacity shift toward assisted review because collaboration is the only feasible way to raise performance without relying extensively on direct human interpretation. The increase in collaboration is therefore sharper for Medicare and Medicaid than for commercial payers. Human engagement remains limited when the required standard is modest and rises only when the performance constraint approaches the human benchmark. This pattern highlights that AI assistance does not automatically imply intensive human oversight. The equilibrium level of oversight depends jointly on reimbursement and on the performance standard imposed on the workflow.