Ask a Buddy or a Bot?
Unbundling Answers from Relationships*

Nicole Immorlica† Brendan Lucier‡

Joseph E. Stiglitz§ Màxim Ventura-Bolet¶

November 3, 2026

This document is a preliminary draft.

Abstract

People routinely turn to other people to answer questions they cannot answer themselves. These consultations deliver answers, but they also build social capital, which raises the quality of future human answers. We develop a dynamic model in which each question can be asked to a human, asked to AI, or left unanswered, and in which the value of asking a human increases with a common stock of social capital that depreciates over time and is replenished by human consultations. AI generates immediate welfare gains: it improves answers to some questions and makes previously unanswered questions worth asking. When it displaces human consultations, however, social capital erodes monotonically and flow welfare declines toward the highest steady state below the pre-AI level, which may involve the collapse of human consultations. The total number of questions asked rises on impact but can eventually fall below its pre-AI level. If long-run welfare falls below the pre-AI benchmark, AI reduces lifetime welfare for sufficiently patient societies. Because individuals do not internalize their contribution to social capital, decentralized users undertake fewer human consultations than a social planner.

*Ventura-Bolet gratefully acknowledges financial support from the Bank of Spain through a doctoral fellowship.

†Microsoft Research and Yale University. Email: nicimm@microsoft.com.

‡Microsoft Research. Email: brlucier@microsoft.com.

§Columbia University and NBER. Email: jes322@columbia.edu.

¶Columbia University. Email: mv3010@columbia.edu.

1 Introduction

People have always systematically turned to other people to answer questions that they themselves can't answer. We can view this as core to our societal information system. "Asking" is a bundled transaction: the same interaction can solve a problem, train a junior colleague, maintain a friendship, transmit tacit context, and connect previously separate parts of a social network. How we acquire and share information has changed over time, with the interdependent advances in technology (printing, social media platforms) and changes in societal interactions. AI represents a major technological advance, a likely transformative moment affecting simultaneously our society and the information ecosystem. The worry is that in spite of its increasing our ability to process and target information, our information ecosystem and therefore societal productivity may deteriorate. Stiglitz and Ventura-Bolet (2026) showed that that might be the case, especially in the short to medium run, because AI and the social platforms disincentivized the gathering and processing of new information, essential inputs into making AI effective. These adverse effects, they suggested, might be particularly relevant in certain domains of knowledge, e.g. in ever evolving economic systems. This paper explores another dimension of human-AI interactions, with similar concerns about adverse consequences, but this time, not just for our information ecosystem, but for societal interactions—what economists sometimes refer to as society's social capital. For the earlier cited problem, there are policy remedies which at least ameliorate the adverse effects (Stiglitz and Ventura-Bolet, 2025). In this case, the remedies are not so apparent.

AI is a highly specialized answer technology. For low-context questions it is well trained, faster, and often more precise than any human whom one might ask; but it lacks context – it does not know the full story behind the question – so for context-heavy questions a human may remain better. And AI unbundles the transaction: it produces answers without producing any additional benefit. The value of asking a human, moreover, rises with the strength of the network – as you and others keep asking, the people asked stay practiced, context stays shared, and ties stay warm – so everyone's consultations feed a common stock of social capital. Each individual, comparing values of the answers question by question, internalizes the worth of her relationships but not the value her consultations create for the people on the other end of them, for the friends who meet through her, or for the wider network. These are akin to but different from the network externalities that arise in financial networks, where firms may take account of the impact of their financial structure including any particular transaction on the terms they face with those with whom they directly interact, but not on the whole financial network.[1]

The equilibrium therefore features *overuse of AI relative to the social optimum*: as the AI improves, decentralized players switch too early to asking it, and we show that it is possible that a better AI can make aggregate welfare decline, even though every choice is a best response to what others are doing and to the environment and no behavioral assumption about irrationality is involved. Whether society and the economy is better off depends on balancing the direct benefits of the superior "answering system" that AI provides vs. the decreased social capital, and correspondingly, the worse answers we get when we decide it is preferable to turn to a human to answer a question, because we know, in spite of AI's strength, the human is likely to provide a better answer. The central object is thus not the quality of AI in isolation but the performance of an ecology containing both humans and AI. We are able to show that while the arrival of AI may improve short term social welfare, as social capital deteriorates, the (flow)

[1] See, e.g., Allen and Gale (2000) and Greenwald and Stiglitz (2003).

social welfare decreases, the long run steady state welfare can be lower than before AI, and with sufficiently high discount factors, lifetime social welfare decreases. Of course, as we show, even before the arrival of AI, the presence of network externalities meant that the system did not maximize societal welfare—there are instances in which from a social point of view, it would be desirable to ask a human (taking into account the effects on social capital), but from a private point of view, it pays simply to leave the question unasked and unanswered. But the arrival of AI can make matters worse: social capital deteriorates, with economic and social consequences, including a poorer information ecosystem, with consequentially poorer economic outcomes.

A theoretical model is useful here because the available evidence is limited in two ways. Earlier technology shocks to how people find answers, such as the arrival of search engines, may not be extrapolable to AI: unlike a search engine, AI answers personalized questions and sustains follow-up conversation, much as a fellow human does. Current empirical evidence, meanwhile, captures only the short run, before the human answering system adjusts and before the equilibrium effects that arise as social capital responds. A theoretical model allows us to isolate the channels at work and trace their consequences over time.

Our paper contributes to a growing theoretical literature on how AI reshapes the production and transmission of knowledge. Ide and Talamàs (2025) study how AI reorganizes knowledge hierarchies. Ide (2026) studies how automation disrupts the transmission of tacit knowledge from experts to novices. Caosun and Aral (2026) analyze the tradeoff between AI's immediate productivity gains and the erosion of worker expertise through cognitive offloading. Acemoglu et al. (2026) show how AI can discourage human learning and lead to a collapse of collective knowledge. We instead study how heterogeneous questions are routed between humans, AI, and leaving them unanswered, and how these decisions change the future value of asking a human. Unlike the information-input channel of Stiglitz and Ventura-Bolet (2026), AI's performance is fixed in our model: potential welfare losses arise solely from the deterioration of the human alternative, which spills over to questions that AI handles poorly.

Authors' statement.

This draft is work in progress. The baseline model treats the answering system in reduced form: social capital is a scalar stock, human consultations feed it through the investment function $I(\cdot)$, it depreciates at rate δ, and the heterogeneity of questions along m is taken as given, without committing to its source. Our research agenda is to microfound this structure with an explicit social network and to ask two questions.

First, *which answering systems are robust to AI?* The dynamics above depend on a handful of reduced-form objects: the depreciation rate, the shape of the investment function (in particular whether x enters I linearly, concavely, or convexly), and the initial conditions when AI arrives, which jointly determine whether the economy lands at a high steady state, a low one, or collapses. We want to understand which properties of an underlying network map into which values of these objects, and hence which answering systems and societies are more or less exposed to the erosion mechanism described here.

Second, *where does question heterogeneity come from?* A natural microfoundation is that answering a question requires two kinds of knowledge: knowledge of the subject, and knowledge of the context and preferences of the person asking. This creates a tradeoff between asking an

expert who knows the subject well but knows you less, and asking a close contact who knows you well but is less expert. Different groups of people are better positioned to answer different subjects and tend to know each other within rather than across groups. We are interested in unbderstanding how AI's arrival should reshape not only how many questions go to humans but which ties are maintained and which parts of the network erode.

2 Pre-AI environment

Time is discrete, $t = 0, 1, 2, \ldots$ There is a mass M of atomistic players. In every period, each player receives one question, indexed by $m \in [0, M]$. They can consult another human or not. If they consult another human, they receive a payoff of $u^H(m, S_t)$ where S_t is the stock of social capital (to be discussed later). If they do not, they receive a payoff of $u^N(m)$.[2]

Define $G(m, S_t) \equiv u^H(m, S_t) - u^N(m)$. We order players/questions so that G is non-increasing in m. That is, as m grows, the gain from consulting a human becomes weakly smaller. To simplify the analysis, we further assume that G is smooth and bounded with $G_m < 0$, $G_S > 0$ and $G_{mS} \geq 0$. The signs of these derivatives have the following meaning: i) $G_m < 0$ makes our ordering assumption more analytically tractable, ii) $G_S > 0$ recognizes that social capital improves the quality of the answers a human can give, iii) $G_{mS} \geq 0$ is a condition that ensures that changes in S do not change the ordering of player/questions. All payoffs are measurable and integrable in m.

Social capital is the value of the structure and quality of social and professional networks. It grows as more human interactions are made. In particular, we assume that the law of motion of social capital is given by:

$$S_{t+1} = (1 - \delta)S_t + I(x_t) \tag{1}$$

where $\delta \in (0, 1)$ is a depreciation rate and $x_t \in [0, M]$ is the measure of questions that are consulted with humans in period t. The function $I : [0, M] \to \mathbb{R}_+$ denotes the increase in social capital that arises when x_t questions are made. We assume that this function is smooth and bounded with $I_x > 0$ for all $x_t \in [0, M]$. That is, human consultations raise social capital.

Since each player is atomistic, her routing choice has no effect on the future stock so they rationally do not consider it and just maximize their utility period by period.[3] This implies that:

$$x_t = x(S_t) = \int_0^M \mathbf{1}\{G(m, S_t) \geq 0\}\, dm \tag{2}$$

where integration is with respect to Lebesgue measure on $[0, M]$. Thus, x_t is simply the measure of the set of questions for which consulting a human is optimal, $\{m \in [0, M] : G(m, S_t) \geq 0\}$. Since the questions are ordered, the following result holds.

Proposition 1 (Measure of human interactions)**.** *The measure of questions asked to humans is given by a threshold:*

[2] We interpret this as leaving the question unanswered but it could also be that the player answers the question on their own.

[3] In a more general framework one could allow each player to have an individual stock of social capital, $S_{m,t}$, and they have an effect on their own network. In this case, each player would face a dynamic problem.

$$x(S_t) = \begin{cases} 0 & \text{if } G(0, S_t) \leq 0 \\ x_t^* & \text{if } G(M, S_t) < 0 < G(0, S_t) \\ M & \text{if } 0 \leq G(M, S_t) \end{cases}$$

where $x_t^* \in (0, M)$ *is implicitly defined by* $G(x_t^*, S_t) = 0$.

As Figure 1 shows, if $G(0, S_t) \leq 0$, then not consulting with humans always provides a higher payoff. If $0 < G(M, S_t)$, then consulting with humans always provides a higher payoff. If $G(M, S_t) < 0 < G(0, S_t)$, our assumptions on G ensure that there exists a unique interior x_t^*.

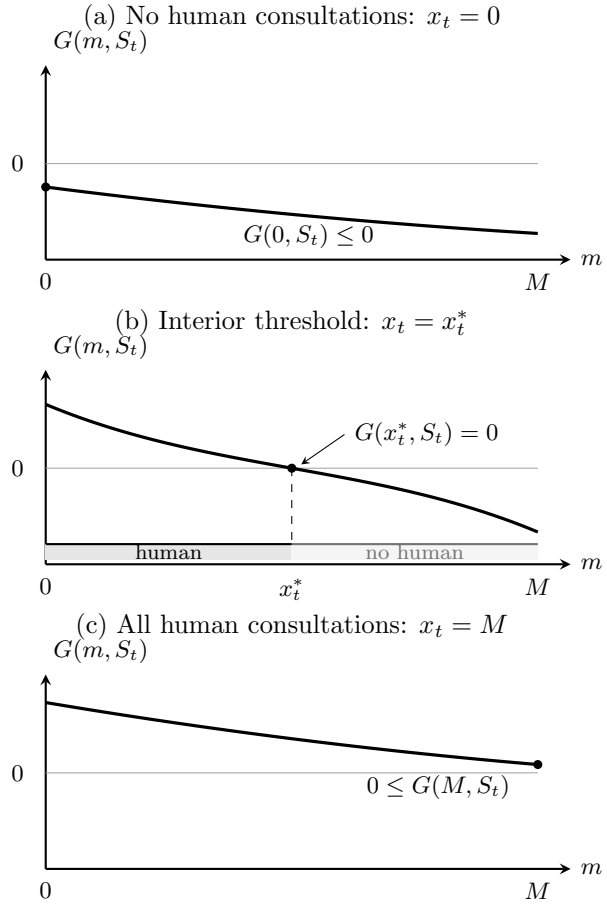


Figure 1: The three possible human-consultation regimes.

The equilibrium of this economy is a sequence $\{x_t, S_t\}_{t=0}^{\infty}$ starting from an initial condition S_0 that satisfies Equation (1) and Equation (2).

Let us define flow welfare as:

$$W(S_t) \equiv \int_{G(m,S_t)\geq 0} u^H(m, S_t)\, dm + \int_{G(m,S_t)<0} u^N(m)\, dm \tag{3}$$

And lifetime welfare as:

$$V(\{S_t\}_{t=0}^{\infty}) \equiv \sum_{t=0}^{\infty} \beta^t W(S_t) \tag{4}$$

where $\beta \in (0,1)$ is the social discount factor.

Notice that flow welfare is increasing in social capital. This allows us to prove that:

Proposition 2. *Consider two sequences $\{S_t\}_{t=0}^{\infty}$ and $\{S'_t\}_{t=0}^{\infty}$. If $S_t > S'_t$ for all t, then $V(\{S_t\}_{t=0}^{\infty}) \geq V(\{S'_t\}_{t=0}^{\infty})$, with strict inequality if $x(S_t) > 0$ for some t.*

A corollary of this result is that the flow welfare at steady state will be highest at whichever steady state has the highest social capital.

From now on we assume that the pre-AI economy is at a steady state, S^{pre-AI}, where some questions are asked to humans and some are not. The following proposition provides a sufficient condition for the steady state to be interior.

Proposition 3 (Pre-AI steady states)**.** *Suppose that $I(0) > 0$, $G(0, I(0)/\delta) > 0$ and $G(M, I(M)/\delta) < 0$, and that every steady state is nondegenerate, i.e., $\frac{d}{dS}\left[I(x(S)) - \delta S\right] \neq 0$ at each steady state. Then there exists an odd number of steady states, all of them interior, i.e, $x_\infty \in (0, M)$. Let n be the number of steady states, then $\frac{n-1}{2} + 1$ of them are stable and the rest are unstable.*

This system always has at least one steady state and possibly more and there are no corner solutions. This last result seems empirically reasonable, i.e, some questions are asked and some are left unanswered.

Figure 2 illustrates the results of this proposition. The top panel shows the case $n = 1$, while the bottom panel shows the case $n = 3$. The condition $I(0) > 0$ rules out a steady state where no questions are asked (which requires that the law of motion starts above zero). The condition $G(M, I(M)/\delta) < 0$ rules out a steady state where all questions are asked (which requires that the law of motion be below the 45° line when $G(M, S) = 0$). Note that Proposition 2 allows us to rank the level of flow welfare in each of these different steady states by social capital.

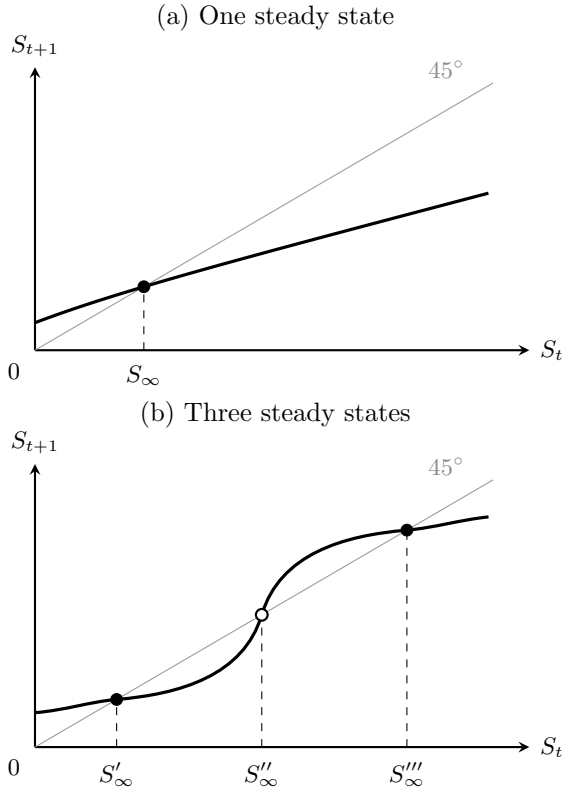


Figure 2: Steady states of the pre-AI social-capital dynamics. Filled circles denote stable steady states; the open circle denotes an unstable steady state.

3 Impact of AI

How does the appearance of AI change things? Suppose that AI becomes available at $t = 0$, when the economy is at its pre-AI steady state, S^{pre-AI}, so that $S_0 = S^{pre-AI}$. Players now have the additional option of asking AI for an answer, yielding a payoff $u^{AI}(m)$. Define:

$$G^{AI}(m, S_t) \equiv u^H(m, S_t) - \max\{u^N(m), u^{AI}(m)\}.$$

This is the gain from consulting a human relative to the best alternative available after AI appears. In particular,

$$\begin{aligned} G^{AI}(m, S_t) &= G(m, S_t) - \max\{0, u^{AI}(m) - u^N(m)\} \\ &\leq G(m, S_t). \end{aligned}$$

At any given stock of social capital, AI therefore weakly reduces the relative attractiveness of consulting a human.

Recall that the pre-AI economy was described by equations 1 - 4. Equation (1) remains the same. But we now have to replace Equation (2), Equation (3) and Equation (4) with:

$$x_t = x^{AI}(S_t) = \int_0^M \mathbf{1}\{G^{AI}(m, S_t) \geq 0\}\, dm \tag{2'}$$

$$W^{AI}(S_t) \equiv \int_{G^{AI}(m,S_t)\geq 0} u^H(m,S_t)\,dm + \int_{G^{AI}(m,S_t)<0} \max\{u^N(m), u^{AI}(m)\}\,dm \tag{3'}$$

$$V^{AI}(\{S_t\}_{t=0}^{\infty}) \equiv \sum_{t=0}^{\infty} \beta^t\, W^{AI}(S_t) \tag{4'}$$

The equilibrium equations have the same structure as before, except that players now have the additional option of consulting AI. Accordingly, G is replaced by G^{AI}, and the payoff from not consulting a human becomes $\max\{u^N(m), u^{AI}(m)\}$. However, G^{AI} need not be decreasing in m, so human consultations need not be characterized by a single cutoff as Proposition 1 established for the pre-AI economy.

3.1 Short Run

Recall that we assume that the pre-AI economy is at a steady state, S^{pre-AI}, where some questions are asked to humans and some are not. When AI appears, it can answer questions that were previously left unanswered and improve answers to questions previously asked to humans. This brings immediate gains to society:

Proposition 4 (Short term gains)**.** *The total increase in flow welfare on impact can be decomposed as:*

$$W^{AI}(S^{pre-AI}) - W(S^{pre-AI}) = \underbrace{\int_{G(m,S^{pre-AI})\geq 0} \max\{0, u^{AI}(m) - u^H(m,S^{pre-AI})\}\,dm}_{\substack{\textit{Better answers to}\\ \textit{previously answered questions}}} + \underbrace{\int_{G(m,S^{pre-AI})<0} \max\{0, u^{AI}(m) - u^N(m)\}\,dm}_{\substack{\textit{Answers to}\\ \textit{previously unanswered questions}}} \geq 0.$$

Both terms are nonnegative. The first is associated with the displacement of human consultations; the second does not directly change human interaction. To gain intuition, we now explore both terms separately through two examples. We first consider the case in which AI answers previously unanswered questions without displacing human consultations.

Example 1 (AI answers previously unanswered questions)**.** *Suppose that AI does not outperform humans on any question previously asked to a human:*

$$u^{AI}(m) \leq u^H(m,S^{pre-AI})$$
$$\textit{for all } m \textit{ such that } G(m,S^{pre-AI}) \geq 0$$

Assume that ties are resolved in favor of consulting a human. Questions in $\{m \in [0,M] : G(m,S^{pre-AI}) \geq 0\}$ *then continue to be asked to humans. For questions in* $\{m \in [0,M] : G(m,S^{pre-AI}) < 0\}$*, players use AI whenever* $u^{AI}(m) > u^N(m)$ *and otherwise retain their*

previous option. AI can therefore increase the number of questions answered without reducing human interaction.

Figure 3 illustrates this case. The solid curve is the pre-AI gain from consulting a human; the dashed curve is the post-AI gain. The curves coincide for $\{m \in [0, M] : G(m, S^{pre-AI}) \geq 0\}$ *in this illustration. Both curves cross zero at the same threshold. To the right of this threshold, the distance between* G *and* G^{AI} *is the payoff gain from having access to AI.*

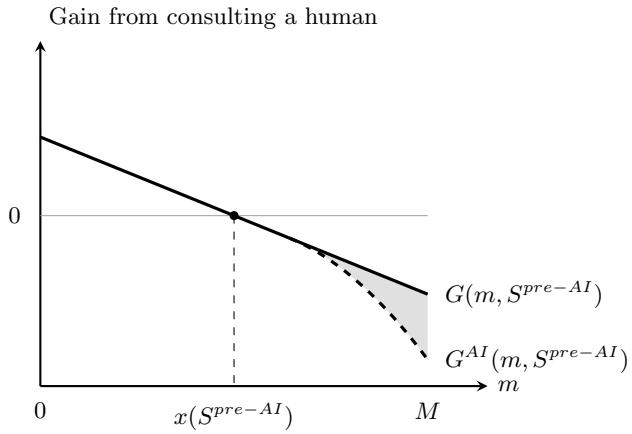


Figure 3: AI answers some previously unanswered questions without displacing human consultations.

Since human consultations remain unchanged, the law of motion implies:

$$S_1 = (1-\delta)S^{pre-AI} + I(x^{AI}(S^{pre-AI})) = S^{pre-AI}$$

AI has no effect on social capital in this case, because it does not change the amount of human interaction. The welfare gains (shaded grey area in Figure 3) are permanent.

AI can also provide better answers to questions that were previously asked to humans. We now isolate this second type of gain.

Example 2 (AI answers substitute for human answers)**.** *Suppose that AI does not improve on the outside option for previously unanswered questions:*

$$u^{AI}(m) \leq u^{N}(m)$$
$$\text{for all } m \text{ such that } G(m, S^{pre-AI}) < 0.$$

However, AI outperforms humans on some questions in $\{m \in [0, M] : G(m, S^{pre-AI}) \geq 0\}$. *Figure 4 illustrates a case in which this occurs on an interval* (m_L^*, m_H^*), *where:*

$$0 < m_L^* < m_H^* < x(S^{pre-AI}).$$

After AI appears, players consult humans for questions in $[0, m_L^*] \cup [m_H^*, x(S^{pre-AI})]$ *and use AI for questions in* (m_L^*, m_H^*). *Questions above* $x(S^{pre-AI})$ *remain unanswered. Thus, the measure of human consultations on impact is:*

$$x^{AI}(S^{pre-AI}) = m_L^* + (x(S^{pre-AI}) - m_H^*) < x(S^{pre-AI}).$$

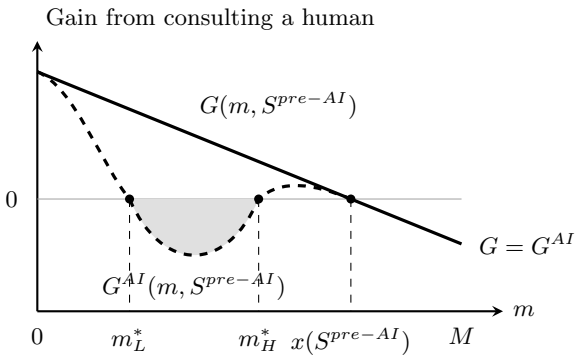


Figure 4: AI improves answers to previously answered questions.

Unlike the pre-AI case, human consultations need not be characterized by a single threshold. Our ordering assumption on G does not imply that G^{AI} is decreasing in m: AI may be particularly effective at answering some questions and less effective at answering others.
The increase in flow welfare is given by the shaded area in Figure 4. Unlike the previous case, however, it is accompanied by a reduction in human interaction. Since $I(\cdot)$ is strictly increasing:

$$S_1 - S^{pre-AI} = I(x^{AI}(S^{pre-AI})) - I(x(S^{pre-AI})) < 0.$$

We assume that both types of initial gains occur and study what happens as social capital adjusts as a result.

3.2 Medium Run

We now explore the dynamics that occur after impact. The decline in social capital changes the quality of human answers and, consequently, subsequent routing choices and welfare.

Proposition 5 (Social capital and human consultations)**.** *The measure of human consultations, $x^{AI}(S_t)$, is nondecreasing in social capital. In particular, for any $S' \leq S$, $x^{AI}(S') \leq x^{AI}(S)$. The inequality is strict whenever the set $\{m \in [0, M] : G^{AI}(m, S') < 0 \leq G^{AI}(m, S)\}$ has positive measure.*

Human consultations may occur on several separate intervals; a decline in social capital weakly shrinks their union. Since AI initially displaces some human consultations, $x^{AI}(S^{pre-AI}) < x(S^{pre-AI})$. Starting from the pre-AI steady state, we therefore have:

$$S_1 - S_0 = I(x^{AI}(S^{pre-AI})) - I(x(S^{pre-AI})) < 0.$$

Proposition 5 then implies $x_1 \leq x_0 = x^{AI}(S^{pre-AI})$. Moreover,

$$S_2 - S_1 = (1 - \delta)(S_1 - S_0) + I(x_1) - I(x_0) < 0.$$

Repeating this argument shows that social capital continues to fall, while human consultations weakly decrease. The feedback loop is illustrated in Figure 5. As S_t falls, $G^{AI}(m, S_t)$ shifts downward, and questions previously asked to humans may instead be directed to AI or left unanswered.

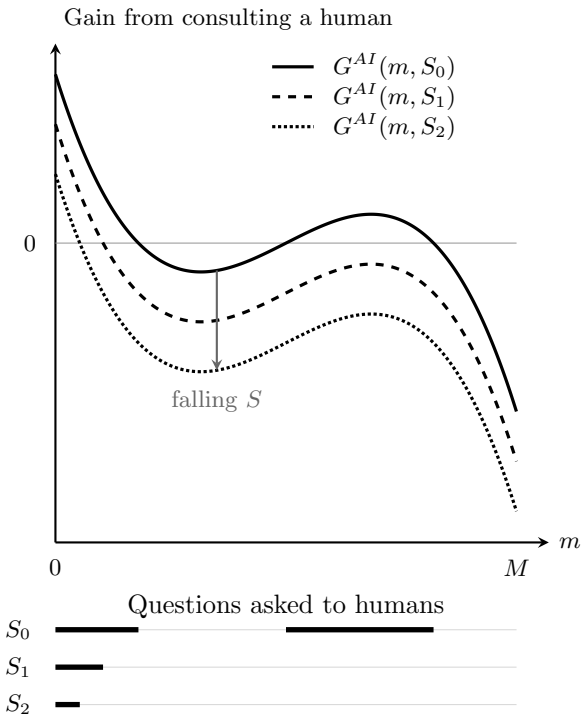


Figure 5: Social capital erosion reduces human consultations.

Proposition 6 (Monotone erosion)**.** *Suppose that AI initially displaces a positive measure of human consultations, so that* $x^{AI}(S^{pre-AI}) < x(S^{pre-AI})$*. Then:*

1. $S_{t+1} < S_t$ *and* $x_{t+1} \leq x_t$ *for every* $t \geq 0$*.*
2. *Social capital converges to a limit*

$$S_\infty^{AI} \in [I(0)/\delta,\, S^{pre-AI}).$$

Proposition 4 establishes that welfare weakly increases on impact, with a strict increase whenever AI strictly improves payoffs on a positive measure of questions. The welfare result for the medium run is not as positive.

Corollary 1 (Declining flow welfare)**.** *Suppose that AI initially displaces a positive measure of human consultations. Then* $W^{AI}(S_{t+1}) \leq W^{AI}(S_t)$ *for every* $t \geq 0$*. The inequality is strict if and only if* $\{m \in [0, M] : G^{AI}(m, S_t) > 0\}$ *has positive measure.*

3.3 Long Run

We next characterize the steady state reached by social capital's negative feedback loop.

Proposition 7 (AI steady-state)**.** *Suppose that AI initially displaces a positive measure of human consultations, so that* $x^{AI}(S^{pre-AI}) < x(S^{pre-AI})$*. Then, the set* $\{S \in [0, S^{pre-AI}] : \delta S = I(x^{AI}(S))\}$ *is nonempty and the economy converges to its largest element. Moreover,*

$$\lim_{t\to\infty} x_t = x^{AI}(S_\infty^{AI}), \qquad \delta S_\infty^{AI} = I(x^{AI}(S_\infty^{AI})).$$

Hence, the stock of social capital starts at S^{pre-AI} when AI arrives and then falls monotonically as more and more questions are displaced away from humans. This feedback loop stops once the stock of social capital reaches the highest steady state below S^{pre-AI}.

We are interested in when the feedback loop leads to a social collapse where no human consultations are made, i.e, $x^{AI}(S^{AI}_{\infty}) = 0$. In this case, the steady state of social capital is given by $S^{AI}_{\infty} \equiv \frac{I(0)}{\delta}$. This stock can be positive if social capital also receives investment unrelated to the consultations modeled here.

The next proposition provides a decomposition of the long run welfare effects.

Proposition 8 (Long run welfare decomposition)**.** *Suppose that AI initially displaces a positive measure of human consultations. Steady state flow welfare is:*

$$\lim_{t\to\infty} W^{AI}(S_t) = \int_0^M \max\{u^H(m, S^{AI}_{\infty}), u^{AI}(m), u^N(m)\}\, dm.$$

Moreover,

$$W^{AI}(S^{AI}_{\infty}) - W(S^{pre-AI}) = \underbrace{W^{AI}(S^{pre-AI}) - W(S^{pre-AI})}_{\textit{Impact gain}} - \underbrace{\int_{\mathcal{D}} G^{AI}(m, S^{pre-AI})\, dm}_{\substack{\textit{Loss on questions} \\ \textit{diverted from humans}}}$$

$$- \underbrace{\int_{\mathcal{K}} \left[u^H(m, S^{pre-AI}) - u^H(m, S^{AI}_{\infty})\right] dm}_{\substack{\textit{Loss on questions} \\ \textit{still asked to humans}}},$$

where

$$\mathcal{D} \equiv \{m : G^{AI}(m, S^{pre-AI}) \geq 0,\ G^{AI}(m, S^{AI}_{\infty}) < 0\},$$
$$\mathcal{K} \equiv \{m : G^{AI}(m, S^{pre-AI}) \geq 0,\ G^{AI}(m, S^{AI}_{\infty}) \geq 0\}.$$

Long-run welfare is below the pre-AI benchmark if and only if the sum of the two erosion losses exceeds the impact gain. The impact gain includes both better answers to previously answered questions and answers to previously unanswered questions, as established in Proposition 4. These gains persist because the nonhuman options remain available. Against them, erosion lowers payoffs on questions that initially remain with humans: some are eventually directed to AI or left unanswered, while others continue to receive human answers from a weaker network.

Besides welfare, one might also be interested in whether the total number of questions asked in society improves thanks to AI. AI expands the set of questions that can be answered without human consultation. However, the erosion of social capital can make other questions no longer worthwhile to ask and AI may be bad at answering those other questions.

Define the set of questions for which AI is at least as good as leaving the question unanswered:

$$\mathcal{A} \equiv \{m \in [0, M] : u^{AI}(m) \geq u^N(m)\}.$$

We retain the rule that ties are resolved in favor of humans and, when humans are not chosen, resolve ties between AI and leaving a question unanswered in favor of AI. This set does not depend on social capital as neither technology uses it, so these are questions that will always be asked in the post-AI economy.

Let Q_t^{AI} denote the total measure of questions asked to either humans or AI at time t. Then:

$$\begin{aligned} Q_t^{AI} &\equiv \int_0^M \mathbf{1}\{\max\{u^H(m, S_t), u^{AI}(m)\} \geq u^N(m)\}\, dm \\ &= |\mathcal{A}| + \int_{[0,M]\setminus\mathcal{A}} \mathbf{1}\{G(m, S_t) \geq 0\}\, dm, \end{aligned}$$

Questions in $\mathcal{A}$ are always answered, either by humans or by AI. Questions outside $\mathcal{A}$ are answered only when consulting a human is preferable to leaving the question unanswered. Before AI arrives, the total measure of questions asked is $Q^{pre-AI} = x(S^{pre-AI})$.

At every date,

$$Q_t^{AI} - x(S^{pre-AI}) = \underbrace{|\mathcal{A} \cap (x(S^{pre-AI}), M]|}_{\substack{\text{Previously unanswered} \\ \text{questions now answered}}} - \underbrace{|\mathcal{B}_t|}_{\substack{\text{Previously answered} \\ \text{questions now unanswered}}},$$

where

$$\mathcal{B}_t \equiv \{m \in [0, x(S^{pre-AI})] : u^{AI}(m) < u^N(m),\ G(m, S_t) < 0\}.$$

A few results are immediate.

Proposition 9 (Total questions asked)**.** *Starting from the pre-AI steady state:*

1. *Total questions asked weakly increase on impact:* $Q_0^{AI} - x(S^{pre-AI}) \geq 0$.
2. *Along the post-AI adjustment path,* $Q_{t+1}^{AI} \leq Q_t^{AI}$ *for every* $t \geq 0$.

   The inequality is strict if and only if the set $\{m \notin \mathcal{A} : G(m, S_{t+1}) < 0 \leq G(m, S_t)\}$ *has positive measure.*
3. *Total questions asked fall below their pre-AI level if and only if*

$$|\mathcal{B}_t| > |\mathcal{A} \cap (x(S^{pre-AI}), M]|.$$

The long-run comparison therefore depends on whether newly answered questions outnumber previously answered questions that become unanswered. Since $Q_t^{AI} \geq |\mathcal{A}|$, total questions asked cannot fall below the pre-AI level if $|\mathcal{A}| \geq x(S^{pre-AI})$. If instead $|\mathcal{A}| < x(S^{pre-AI})$ and human consultations collapse, then $Q_\infty^{AI} = |\mathcal{A}| < x(S^{pre-AI})$. Thus, an initial expansion in total questions asked can eventually be reversed.

4 Social Planner

An important observation is that the pre-AI world might already exhibit an inefficient social capital level due to the existence of an externality. Each atomistic player is not taking into account their effect on the future stock of social capital but this improves the payoff everyone obtains from asking humans questions in the future. This might lead to questions where it would be socially optimal to ask a human but are not asked because it is not privately optimal.

First we establish that if steady state flow welfare is lower than in the pre-AI economy, lifetime welfare falls for sufficiently patient societies.

Proposition 10 (Lifetime welfare and patience)**.** *Suppose that AI initially displaces a positive measure of human consultations. Let $\{S_t\}_{t=0}^{\infty}$ be the social capital equilibrium dynamics under AI where $S_0 = S^{pre-AI}$. Then:*

1. *If $W^{AI}(S_\infty^{AI}) \geq W(S^{pre-AI})$, then*

$$V^{AI}(\{S_t\}_{t=0}^{\infty}) > \frac{W(S^{pre-AI})}{1-\beta} \quad \text{for all } \beta \in (0,1).$$

2. *If $W^{AI}(S_\infty^{AI}) < W(S^{pre-AI})$, there exists a unique $\beta^* \in (0,1)$ such that*

$$V^{AI}(\{S_t\}_{t=0}^{\infty}) - \frac{W(S^{pre-AI})}{1-\beta} \begin{cases} > 0, & \beta < \beta^*, \\ = 0, & \beta = \beta^*, \\ < 0, & \beta > \beta^*. \end{cases}$$

If long-run welfare remains above the pre-AI benchmark, the initial gains are worthwhile for every discount factor. If it falls below that benchmark, sufficiently patient societies place enough weight on the subsequent losses hence AI reduces lifetime welfare.

The Social Planner maximizes the present discounted value of society's welfare. The planner therefore chooses which questions are asked to humans in each period, taking into account their contribution to future social capital. Questions not assigned to humans receive their best nonhuman option.

Let $\mathcal{H}_t \subseteq [0, M]$ denote the measurable set of questions assigned to humans. The planner solves:

$$\begin{aligned} \max_{\{\mathcal{H}_t\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t \left[\int_{\mathcal{H}_t} u^H(m, S_t)\, dm + \int_{[0,M]\setminus\mathcal{H}_t} \max\{u^N(m), u^{AI}(m)\}\, dm \right] \\ \text{s.t.} \quad & S_{t+1} = (1-\delta)S_t + I(|\mathcal{H}_t|), \qquad S_0 = S^{pre-AI}. \end{aligned} \tag{5}$$

Here $|\cdot|$ denotes Lebesgue measure, so the measure of human consultations is $x_t = |\mathcal{H}_t|$. Conditional on the measure of human consultations, the planner assigns to humans the questions with the highest values of $G^{AI}(m, S_t)$. These questions need not form a single interval.

The decentralized allocation maximizes current flow welfare. The planner additionally values the future social capital generated by human consultations and may therefore assign questions to humans even when a nonhuman option offers a higher current payoff. Among equally optimal allocations, we select one with the largest measure of human consultations.

Proposition 11 (Insufficient decentralized human consultations)**.** *Let $\mathcal{H}^P(S)$ denote the planner's selected optimal consultation set at stock S. Then*

$$|\mathcal{H}^P(S)| \geq x^{AI}(S).$$

Starting from the same stock S^{pre-AI}, the planner's path satisfies

$$S_t^P \geq S_t, \qquad |\mathcal{H}^P(S_t^P)| \geq x^{AI}(S_t) \qquad \text{for every } t \geq 0.$$

Decentralized users choose the best current answer but do not internalize how their consultations improve future human answers. The planner preserves weakly more human interaction and social capital, potentially sacrificing current payoff to improve future welfare.

5 Conclusion

AI can improve answers and make previously unanswered questions worth asking, generating immediate welfare gains. Yet when it displaces human consultations, the resulting erosion of social capital can reduce the quality of human answers and leave sufficiently patient societies worse off. Evaluating AI adoption therefore requires accounting for both its direct benefits and its effects on the human relationships that support future information exchange.

References

Daron Acemoglu, Dingwen Kong, and Asuman Ozdaglar. AI, human cognition and knowledge collapse. NBER Working Paper 34910, National Bureau of Economic Research, February 2026.

Franklin Allen and Douglas Gale. Financial contagion. *Journal of Political Economy*, 108(1): 1–33, 2000.

Michael Caosun and Sinan Aral. The augmentation trap: AI productivity and the cost of cognitive offloading. arXiv:2604.03501, 2026.

Bruce Greenwald and Joseph E. Stiglitz. *Towards a New Paradigm in Monetary Economics*. Cambridge University Press, Cambridge, 2003.

Enrique Ide. Automation, AI, and the intergenerational transmission of knowledge. arXiv:2507.16078, 2026.

Enrique Ide and Eduard Talamàs. Artificial intelligence in the knowledge economy. *Journal of Political Economy*, 133(12):3762–3800, 2025. doi: 10.1086/737233.

Joseph E. Stiglitz and Màxim Ventura-Bolet. Information in the age of AI: Challenges and solutions. In *The Digitalist Papers, Volume 2: The Economics of Transformative AI*. Stanford Digital Economy Lab, December 2025.

Joseph E. Stiglitz and Màxim Ventura-Bolet. The impact of AI and digital platforms on the information ecosystem. In Ajay K. Agrawal, Erik Brynjolfsson, and Anton Korinek, editors, *The Economics of Transformative AI*, pages 367–389. University of Chicago Press, Chicago, 2026.