Human Judgment and AI Pricing

By AJAY AGRAWAL, JOSHUA S GANS AND AVI GOLDFARB*

* Agrawal, Gans, Goldfarb: Rotman School of Management, University of Toronto, 105 St George St, Toronto ON M5R 2X9 (e-mail: joshua.gans@utoronto.ca) and NBER.

I. Introduction

Artificial intelligence (AI) is undergoing a renaissance. Thanks for developments in machine learning – particularly, deep learning and reinforcement learning – there has been an explosion in the applications of AI in many settings. In actuality, however, far from providing new forms of machine intelligence in a general fashion, what AI has been able to do has been to reduce the cost of higher quality predictions in a drastic way (Agrawal et.al., 2018). As deep learning pioneer Geoffrey Hinton put it, "Take any old problem where you have to predict something and you have a lot of data, and deep learning is probably going to make it work better than the existing techniques." (Hinton 2016) Thus, when they are able to utilize AI, decision-makers know more about their environment including about future states of the world.

These developments have brought about discussion as to the role of humans in that decision-making process. The view we take here (see also Agrawal et.al., 2017) is that humans still play a critical role in determining the reward functions in decisions. That is, if the decision can be formulated as a problem of choosing an action (x), in the face of uncertainty about the state of the world (θ) with probability distribution function $F(\theta)$, in an ideal setting, an AI can transform that problem from $max_x \int u(x,\theta)dF(\theta)$ into $max_x u(x,\theta)$ with actions being made in a state-contingent manner. However, this transformation relies on someone knowing the utility function, $u(x,\theta)$. We argue that, at present, only a human can develop this knowledge.

That said, the value to understanding the utility function in all of its nuance is enhanced when the decision-maker knows that they will have accurate predictions of the state of the world. This is especially true when it comes to states that are unlikely to arise or as applied to decision-making in complex environments.

Here we develop a model of utility function discovery in the presence of AI. In so doing, we choose to emphasize experiences as the means by which decision-makers come to know that function. Our goal here is to understand what this implies for the demand for AI and, in particular, how suppliers of AI services should go about pricing their services. We show that learning leads to some interesting dilemmas in setting AI pricing. In particular, learning may lead decision-makers to discover they have dominant actions and so do not need AI for prediction at all. This presents challenges for the long-term pricing of AI services. The mechanism driving this result is related to the price discrimination literature on the strategic effects of firms gaining information about consumers (e.g. Hart and Tirole 1988; Villas-Boas 2004; Acquisti and Varian 2005; Zhang 2011; Fudenberg and Villas-Boas 2012).

II. Model Set-Up

Our baseline model is drawn from Agrawal et.al. (2017); itself inspired by Bolton and Faure-Grimaud (2009). The decision-maker faces uncertainty over two states of the world, $\{\theta_1, \theta_2\}$ with equal prior probabilities. There are two possible actions: a state independent action with known payoff of S (safe) and a state dependent action with unknown payoff, R or r as the case may be (risky). The agent does not know the payoff from the risky action in each state and must apply *judgment* to determine that payoff. We assume that there are only two possible payoffs from the risky action, R and r, where $R > S > r$. In the absence of judgment, the ex ante expectation that the risky action is optimal in state θ_i is v; common across states. That is, v is the probability in state θ_i that the risky payoff is R rather than r. This is a statement about the payoff, given the state.

In the absence of knowledge regarding the specific payoffs from the risky action, a decision can only be made on the basis of prior probabilities. In this case, the expected payoff from the risky action is $\rho \equiv vR + (1-v)r$. We make the following assumptions:

A1 (Safe Default) $\rho \leq S$

A2 (Judgment Insufficient) $\frac{1}{2}(R + r) \leq S$

(A1) states that, in the absence of judgment, the safe action is the default in each state. (A2) states that, if the agent knows the payoffs in each state, judgment alone will not change that default.

If an AI is deployed to assist in this decision-making, what it does is provide an ex ante prediction of the state. To keep things simple, we assume that prediction is perfect and so, with an AI, the decision-maker knows the state with certainty. By (A2), without judgment, having an AI does not change the decision or payoff. With both judgment and a prediction, optimal state-contingent decision-making is possible and the decision-maker's expected payoff is $\rho^* \equiv vR + (1-v)S$ in each period.

III. Judgment Through Experience

Judgment does not come for free. In Agrawal et.al. (2017), we assume that it takes thought (at the cost of time). By contrast, here we assume that judgment arises from experience. Specifically, an agent must actually experience a given state in order to, potentially, learn the payoffs from that state. If they do not know the state, they cannot learn.

Decision-makers discount with factor $\delta < 1$. If a state arises, there is a probability, λ, that they will gather enough experience to determine the optimal action in that period and can make a choice based on that judgment. Otherwise, they can make a decision in the absence of that judgment. Importantly, they cannot learn the payoff associated with the state if they take the default action. Ignorance remains and their per period payoff is *S*.

The timing of the game is as follows:

1. (Prediction) The decision-maker is informed by the AI of the state that period.
2. (Judgment) With probability, $1-\lambda$, the decision-maker does not learn the payoffs for the risky action in that state. With probability λ, the decision-maker gains this knowledge and retains it into the future.
3. (Action) Based on these outcomes, the decision-maker takes an action and payoffs are realized and we move to the next time period.

There are three phases to experience: (i) *Full experience*: when the agent has learned payoffs in both states, resulting in a discounted payoff from this point of: $\frac{1}{1-\delta}\rho^*$. (ii) *Partial experience*: Let π_i denote the expected present discounted value if the agent already knows what the optimal action is in θ_i. Then: $\pi_i = \frac{1}{2}(\rho^* + \delta\pi_i) + \frac{1}{2}\left((1-\lambda)(S+\delta\pi_i) + \lambda\frac{1}{1-\delta}\rho^*\right)$

$$\Rightarrow \pi_i = \frac{(1+\frac{\lambda}{1-\delta})\rho^* + (1-\lambda)S}{2(1-(1-\frac{1}{2}\lambda)\delta)}$$

And finally, (iii), *no experience* with expected discounted payoff of:

$$\Pi = \lambda(\rho^* + \delta\pi_i) + (1-\lambda)(S+\delta\Pi)$$

$$\Rightarrow \Pi = \frac{(1-\delta)(1-\lambda)2S + (2-\delta)\lambda\rho^*}{(1-\delta)(2-(2-\lambda)\delta}$$

Thus, there is a learning period of uncertain length followed by a period whereby the agent can apply full experience to decisions into the future earning ρ^* on average. As λ increases, so does Π, showing that prediction and judgment are complements in this model.

IV. Pricing AI as a Service

Without any judgment or experience, the net present discounted value earned by the agent would be $\frac{1}{1-\delta}S$. Without initial access to an AI, the agent cannot apply judgment and gain experience to improve upon this. This suggests that a monopolist provider of AI could charge a fixed sum of $\Pi - \frac{1}{1-\delta}S$. Moreover, as Π is increasing in λ, that provider would want to target agents with judgment ability (or ease) as high as possible first before moving on to worse judges.

There are several challenges to pricing AI with a once off payment. First, such algorithms often are run by the provider and not hosted as a distinct app by the user. Therefore, there are on-going costs to be recouped and users may be reluctant to pay up front for such a service. Second, algorithms hosted by the provider may improve at a more rapid rate. The provider may then want a means of monetizing those improvements.

For these reasons, we consider pricing of AI as an ongoing service with a subscription fee of p per period. If the AI provider does not have knowledge of the experience level – and indeed, the experience – of each agent, this is a non-trivial pricing problem.

To see this, let us consider the purchase decisions of fully experienced agents who know their payoff function. For some of these agents, they would have found that neither the safe nor risky action is dominated and their per period expected payoff is $\frac{1}{2}(R+S)$. They can realize these payoffs with prediction but in the absence of prediction, they earn S per period (by A1). Thus, their willingness to pay for prediction is $\frac{1}{2}(R-S)$. For other agents, their experience has shown them that one of the actions is dominated. Those agents either earn R or S per period but do not need prediction to do so. What this means is that the long-term market for prediction is at most a share $2v(1-v)$ of the original market; that is, prediction is only valuable to those who have found neither action to be dominated. To keep things simple, in this section we assume that $v=\frac{1}{2}$. In this case, a fully experienced agent will continue to purchase AI if $\frac{1}{2}(R-S)\geq p$. If the provider, charges a price based on this, they will earn $\frac{1}{4}(R-S)$ per period.

What determines whether a partially experienced agent pays for the AI service? If they have learned that the risky action is optimal in one state, their expected discounted payoff is π_R where:

$$\begin{aligned}
&\pi_R \\
&= \frac{1}{2}(R+\delta\pi_R) \\
&+ \frac{1}{2}\left((1-\lambda)(S+\delta\pi_R)+\lambda\frac{(3R+S-2p)}{4(1-\delta)}\right)-p \\
&\Longrightarrow \pi_R = \frac{R+\frac{\lambda}{4(1-\delta)}(3R+S-2p)+(1-\lambda)S-2p}{2\left(1-\left(1-\frac{1}{2}\lambda\right)\delta\right)}
\end{aligned}$$

If this agent did not have access to an AI after this point, their expected discounted payoff would be: $\frac{1}{1-\delta}\max\{\frac{1}{4}(3R+r),S\}$. On the other hand, if a partially experienced agent learned the safe action was optimal in one state, their expected discounted payoff is π_S where:

$$\pi_S = \frac{1}{2}(S + \delta\pi_S) + \frac{1}{2}\left((1-\lambda)(S+\delta\pi_S) + \lambda\frac{(R+3S-2p)}{4(1-\delta)}\right) - p$$

$$\Longrightarrow \pi_S = \frac{S + \frac{\lambda}{4(1-\delta)}(R+3S-2p) + (1-\lambda)S - 2p}{2\left(1-\left(1-\frac{1}{2}\lambda\right)\delta\right)}$$

If this agent did not have access to an AI after this point, their expected discounted payoff would be: $\frac{1}{1-\delta}S$. These two options differ both in terms of the payoffs they generate while learning as well as what the potential upside is from moving to full experience. If the agent has learned that the risky action is optimal, this upside is $\frac{3}{4}R + \frac{1}{2}S - p$ while otherwise it is $\frac{1}{2}R + \frac{3}{4}S - p$. Thus, $\pi_R > \pi_S$.

This leads to a pricing dilemma on the part of an AI provider. They have two pricing options: they can set p so that $\min\{\pi_R - \frac{1}{1-\delta}\max\left\{\frac{1}{4}(3R+r), S\right\}, \pi_S - \frac{1}{1-\delta}S\} \geq 0$ thereby, selling to the entire market or price above this level so that either $\pi_R \geq \frac{1}{1-\delta}\max\left\{\frac{1}{4}(3R+r), S\right\}$ or $\pi_S \geq \frac{1}{1-\delta}S$ and sell to half of the market. The following proposition demonstrates, however, that, for a far-sighted AI provider, servicing the entire market is the more profitable approach; however, the AI provider does not extract the full value of the prediction despite having perfect knowledge of the state.

Proposition 1. *For δ sufficiently high, the AI provider will maximize profits by covering the entire market with a price equal to:*

$$p = \frac{\lambda}{2(\lambda + 4(1-\delta))}(R - S).$$

PROOF: We first examine the prices in the proposition that would result in full inclusion. The prices are such that $\pi_R \geq \frac{1}{1-\delta} max\{\frac{1}{4}(3R + r), S\}$ and $\pi_S \geq \frac{1}{1-\delta}S$ so there is full inclusion with partially experienced agents. Specifically, note that the price where $\pi_R = \frac{1}{1-\delta}S$, $p = \frac{(4(1-\delta)+3\lambda)}{2(\lambda+4(1-\delta))}(R-S) > \frac{\lambda}{2(\lambda+4(1-\delta))}(R-S)$ which is the price where $\pi_S = \frac{1}{1-\delta}S$. It is useful to examine whether these prices will result in inclusion with fully experienced agents. Note also that the price where $\pi_R = \frac{1}{4(1-\delta)}(3R+r)$, $p = \frac{2(2S-R-r)(1-\delta)+\lambda(R-S+\delta(2S-R-r))}{2(\lambda+4(1-\delta))} > \frac{\lambda}{2(\lambda+4(1-\delta))}(R-S)$. Thus, the price in the proposition is the only price that will support full inclusion at the partially experience phase.

Will this price also support inclusion at the fully experience stage; that is, is $p \leq \frac{1}{2}(R-S)$? Note that $\frac{\lambda}{2(\lambda+4(1-\delta))} < \frac{1}{2}$ so this is satisfied. Note, however, that $\frac{(4(1-\delta)+3\lambda)}{2(\lambda+4(1-\delta))}(R-S) > \frac{1}{2}(R-S)$ and that, as $\delta \to 1$, $\frac{2(2S-R-r)(1-\delta)+\lambda(R-S+\delta(2S-R-r))}{2(\lambda+4(1-\delta))} \to \frac{(R-S+\delta(2S-R-r))}{2} > \frac{1}{2}(R-S)$. Thus, under these conditions, setting a price that excludes some agents at the partial experience phase, causes future demand by fully experienced agents to fall to 0.

When we examine pricing to agents with no experience, note that:

$$\Pi = \lambda\frac{1}{2}(R+\delta\pi_R+S+\delta\pi_S)+(1-\lambda)(S+\delta\Pi)-p \Longrightarrow \Pi = \frac{\lambda\frac{1}{2}(R+\delta\pi_R+S+\delta\pi_S)+(1-\lambda)S-p}{1-(1-\lambda)\delta}$$

The issue is whether an AI provider can charge a price that extracts the maximal rents at this phase; i.e., so that $\Pi = \frac{1}{1-\delta}S$. If this could be done, p will be: $p = \lambda\frac{1}{2}\left(R+\delta\pi_R+S+\delta\pi_S-\frac{2}{1-\delta}S\right)$. If

Substituting and solving for p we have:

$$p = \lambda\frac{(2-\delta)(1-(1-\lambda)\delta)}{4-\delta\left(8-\lambda(6+\lambda)-\delta(4-6\lambda)\right)}(R-S)$$

However, it easy to check that at this price $\pi_S < \frac{1}{1-\delta}S$, so this would not result in full inclusion beyond that phase. Moreover, as $\delta \to 1$, this price becomes $(R-S)$. Therefore, the price in the proposition is the only fully inclusive price resulting in a long-run per period payoff of more than $\frac{1}{2}p = \frac{\lambda}{4(\lambda+4(1-\delta))}(R-S)$ as the provider always serves half of the fully experienced agents.

We have also shown that for δ sufficiently high, any candidate exclusionary price will result it prices that exceed $\frac{1}{2}(R-S)$. Thus, for δ sufficiently high, the AI provider will not find it profitable to exclude agents at any stage. □

Intuitively, when some initial judgment is complete, there is either good news (in that the risky strategy is optimal) or bad news (in which it is not). An inclusion strategy requires price to be low enough that following bad news, learning still occurs. However, while the upside potential for the user following good news is higher than that following bad news, the value of prediction after full experience is gained is the same. Thus, the AI provider has no mechanism by which they can share in the upside. In the absence of that mechanism, they choose to price low and not exclude any users at this stage. Half of the users eventually opt out when they find that either the safe or risky action is dominant. What this means is that an AI provider who cannot implement upfront pricing is restricted in the value they can appropriate. While learning can yield good or bad news to the decision-maker, good news may cause prediction to lose its value as the decision-maker discovers the risky action is dominant. Thus, the AI provider must sacrifice rents in order to ensure that they can capture some rents as the decision-maker gains experience.

Can versioning – selling an AI product which has lower performance – improve this outcome for the AI provider? The intuition would be that until they are fully experienced, users will purchase the lower performing product allowing the AI provider to charge more in the long-term. The downside is a lower performing product may slow the gathering of experience and push that long-term out further. The details of this are left to future work.

V. Judgment Through Experimentation

Using an experience frame to judgment suggests an alternative way of ‘learning’ the reward function: experimentation. In particular, when coupled with prediction, a decision-maker could, by choosing the risky action, evaluate whether that is the right action for that state. The expected cost would be $S - \rho$. In this conception, we have the following:

$$\pi_i = \frac{1}{2}(\rho^* + \delta\pi_i) + \frac{1}{2}\left(\rho + \delta\frac{1}{1-\delta}\rho^*\right)$$

$$\Longrightarrow \pi_i = \frac{\frac{1}{1-\delta}\rho^* + \rho}{2-\delta}$$

Thus, the expected present discount payoff prior to any experience is:

$$\Pi = \rho + \delta\pi_i$$

$$\Longrightarrow \Pi = \frac{1}{2-\delta}\left((3-\delta)\rho + \frac{\delta}{1-\delta}\rho^*\right)$$

The convenient property of this frame is that it relates the cost of judgment explicitly to the expected cost of experimentation. In particular, as r decreases, experimentation becomes more costly.

These calculations presume that the decision-maker finds it worthwhile to experiment. If no experiment is undertaken, the present discounted payoff is $\frac{1}{1-\delta}S$. Thus, it may be the case that there is no value for an AI as the cost of experimentation may be too high.

Even if this were not the case, in assessing the demand for AI under experimentation, we need to consider the fact that decision-makers can use experimentation to discover whether they have dominated actions or not. Depending on v, by running repeated experiments, even in the absence of knowledge of which state has arisen, the decision-maker can potentially infer whether the risky or safe action is preferred in both states. In this case, as we noted earlier, there would be no further demand for an AI.

Working out the full equilibrium outcome under experimentation is beyond the scope of our analysis in this short paper. However, we believe that, in some environments, this could prove to be an interesting driver of the demand for AI and how it evolves.

REFERENCES

Acquisti, A., and H. Varian (2005), “Conditioning Prices on Purchase History,” *Marketing Science* 24(3), 367-381.

Agrawal, A., J.S. Gans and A. Goldfarb (2017), “Prediction Machines, Judgment and

Complexity," in *The Economics of Artificial Intelligence*, A. Agrawal et.al. (eds), University of Chicago Press: Chicago (forthcoming).

Agrawal, A., J.S. Gans and A. Goldfarb (2018), *Prediction Machines: The Simple Economics of Artificial Intelligence*, Harvard Business Review Press: Boston.

Bolton, P. and A. Faure-Grimaud (2009), "Thinking Ahead: The Decision Problem," *Review of Economic Studies*, 76: 1205-1238.

Fudenberg, D., and J. M. Villas-Boas (2012), "Price Discrimination in the Digital Economy." Chapter 10 in *Oxford Handbook of the Digital Economy*, eds. M. Peitz and J. Waldfogel.

Hinton, G, (2016), "Geoff Hinton: On Radiology" Panel discussion, October 27, https://www.youtube.com/watch?v=2HMPRXstSvQ.

Villas-Boas, J. M. (2004), "Price cycles in markets with customer recognition," *RAND Journal of Economics* 35(3), 486-501.

Zhang, Juanjuan (2011), "The Perils of Behavior-Based Personalization," *Marketing Science* 30(1), 170-18.