Delegation of Learning from Multiple Sources of Information

Angelos Diamantopoulos
SOAS, University of London

Arina Nikandrova
City, University of London

CRETE 2023, Naxos
July 14, 2023

Context

- A Principal wants to take an action that matches the unknown state.
- She can hire a biased Expert to learn on her behalf.
- Example: a financial adviser learning about the value of two stocks.

Context

- A Principal wants to take an action that matches the unknown state.
- She can hire a biased Expert to learn on her behalf.
- Example: a financial adviser learning about the value of two stocks.

QUESTION: Would she ever choose to allow the expert to learn and act on her behalf or would she want to retain the option to act also?

Context

- A Principal wants to take an action that matches the unknown state.
- She can hire a biased Expert to learn on her behalf.
- Example: a financial adviser learning about the value of two stocks.

QUESTION: Would she ever choose to allow the expert to learn and act on her behalf or would she want to retain the option to act also?

MAIN RESULT: If the expert's preferences are sufficiently aligned, there are beliefs about the state for which she would prefer to delegate both learning and decision making.

Environment

- Time: $t > 0$
- Discount rate: r
- Two players: Principal P and Expert E
- Two states of the world, $\omega \in \{A, B\}$
- Two actions $x \in \{a, b\}$
- Payoff profiles (u_x^ω, v_x^ω)

P, E	A	B
a	$1,\ 1 + \Delta$	$-1,\ \Delta - 1$
b	$\Delta - 1,\ -1$	$1 + \Delta,\ 1$

- P and E prefer to match the state: $1 > \Delta$
- E is biased Δ

Assumptions

Assumption: $\Delta \leq \max\left\{\bar{\Delta}\left(r\right), 1\right\}$, where $\bar{\Delta}\left(r\right) > 0$

- This restriction ensures that if the expert finds it optimal to learn forever, then the principal also prefers learning to her outside option.
- This restriction has the direct implication that when the expert stops learning under complete delegation, the principal and the expert agree on the optimal decision and thus the principal derives some value from the expert's learning.

Literature

Optimal dynamic learning

- **Che and Mierendorff, 2019: Sequential decision making with two states**
- Zhong, 2022: Optimal learning with flexible signal process and costly information acquisition
- Mayskaya, 2020: Learning with three states
- Nikandrova and Pancs, 2018: Decision problem with 4 states

Delegation of learning and decision making

- Escobar and Zhang, 2021: Optimal truth-revealing contract for disclosure of private learning
- Herresthal, 2022: Trade-off between observable and unobservable learning
- Dessein, 2002: Principal prefers to delegate learning because they cannot elicit truthful information from biased expert
- Argenziano, Severinov, and Squintani, 2016: Delegation versus communication between expert and principal

Complete vs Flexible Delegation

Complete vs Flexible Delegation

Complete Delegation

- P hires E to learn about ω and make a decision $x \in \{a, b\}$.
- Essentially, the E acts according to his preferences as in the decision problem in Che and Mierendorff (2019) and P decides at $t = 0$ whether or not to delegate.

Complete vs Flexible Delegation

Complete Delegation

- P hires E to learn about ω and make a decision $x \in \{a, b\}$.
- Essentially, the E acts according to his preferences as in the decision problem in Che and Mierendorff (2019) and P decides at $t = 0$ whether or not to delegate.

Flexible Delegation

- At $t > 0$, E learns about ω and can make decision $x \in \{a, b\}$
- P can override E and make decision $y \in \{a, b\}$.

Learning

Learning

- Two sources of information: A and B.
- 1 unit of learning intensity, α, to be allocated each dt between the sources.
- If $\omega = A$ and E devotes learning intensity $\alpha \in [0,1]$ to source A for $dt > 0$, *signal* arrives with probability $1 - e^{-\alpha dt}$
- If $\omega = B$ and E devotes learning intensity $1 - \alpha \in [0,1]$ to source B for $dt > 0$, *signal* arrives with probability $1 - e^{-(1-\alpha)dt}$
- $p(signal|\omega = A, \alpha = 0) = p(signal|\omega = B, \alpha = 1) = 0$

Induced belief

- Learning is public $\rightarrow$ belief $p_t \equiv Pr\{\omega = A|$ no signals up to $t\}$ is common
- As long as no signal arrives up to t,

$$\dot{p}_t = -(2\alpha_t - 1)p_t(1 - p_t)$$

Heuristic Time-line

1. E chooses $x_t \in \{\emptyset, a, b\}$
2. If $x_t = \emptyset$. E chooses α_t
3. Having observed α_t, P decides $y_t \in \{\emptyset, a, b\}$
4. If $y_t = \emptyset$, learning outcomes are observed and belief p_t is updated

Markov Strategies

E's strategy

- $\sigma^E = (\alpha, X)$ maps p into decision $X(p) \in \{\emptyset, a, b\}$ and learning intensity $\alpha(p) \in [0, 1]$
- σ^E induces stopping time τ^E when E stops learning

Markov Strategies

E's strategy

- $\sigma^E = (\alpha, X)$ maps p into decision $X(p) \in \{\emptyset, a, b\}$ and learning intensity $\alpha(p) \in [0, 1]$
- σ^E induces stopping time τ^E when E stops learning

P's strategy

- $\sigma^P = Y$ maps p and α into decision $Y(p, \alpha) \in \{\emptyset, a, b\}$
- σ^P induces stopping time τ^P when P stops learning

Markov Strategies

E's strategy

- $\sigma^E = (\alpha, X)$ maps p into decision $X(p) \in \{\emptyset, a, b\}$ and learning intensity $\alpha(p) \in [0, 1]$
- σ^E induces stopping time τ^E when E stops learning

P's strategy

- $\sigma^P = Y$ maps p and α into decision $Y(p, \alpha) \in \{\emptyset, a, b\}$
- σ^P induces stopping time τ^P when P stops learning

$\tau \equiv \tau^E \wedge \tau^P$ is a stopping time when all learning stops
$\{p^\sigma(t) | t \geq 0\}$ is the belief process induced by strategy profile $\sigma = (\sigma^E, \sigma^P)$

Markov Equilibrium

For any σ and initial belief p_0, P's payoff

$$J^P(p_0, \sigma) \equiv \mathbb{E}\Big[\mathbb{1}\{\tau = \tau^E\} e^{-rt} U_{X(\tau)}(p^\sigma(\tau)) \\ + \mathbb{1}\{\tau = \tau^P\} e^{-rt} U_a(p^\sigma(\tau)) \vee U_b(p^\sigma(\tau)) | p_0\Big]$$

For any σ and initial belief p_0, E's payoff

$$J^E(p_0, \sigma) \equiv \mathbb{E}\Big[\mathbb{1}\{\tau = \tau^E\} e^{-rt} (V_\alpha(p^\sigma(\tau)) \vee V_b(p^\sigma(\tau))) \\ + \mathbb{1}\{\tau = \tau^P\} e^{-rt} V_{Y(\tau)}(p^\sigma(\tau)) | p_0\Big]$$

An equilibrium strategy profile $\sigma^* = (\sigma^{E*}, \sigma^{P*})$ for all p satisfies

$$\sup_{\sigma^P} J^P(p, (\sigma^{E*}, \sigma^P)) \text{ and } \sup_{\sigma^E} J^P(p, ((\sigma^E, \sigma^{P*})$$

- We characterize this by "guessing and verifying". HJB

Main Result

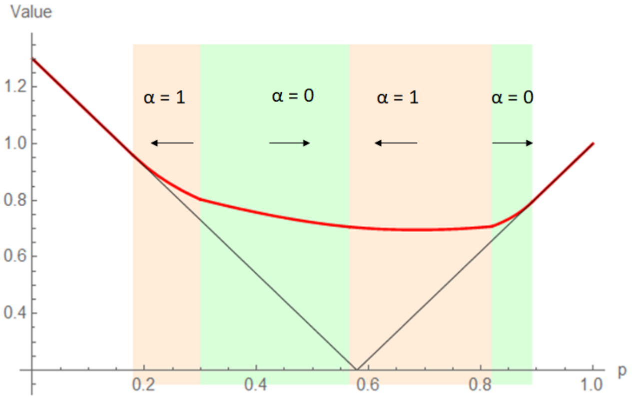
A Markov equilibrium $\sigma^* = (\sigma^{E*}, \sigma^{P*})$ exists.

For any r such that learning is individually optimal for the expert, there exists a $\underline{\Delta}(r)$ such that

- for $\Delta \leq \underline{\Delta}(r)$, P is better off under complete delegation than under flexible delegation for some p
- for $\Delta > \underline{\Delta}(r)$, P is better off under flexible delegation for any p

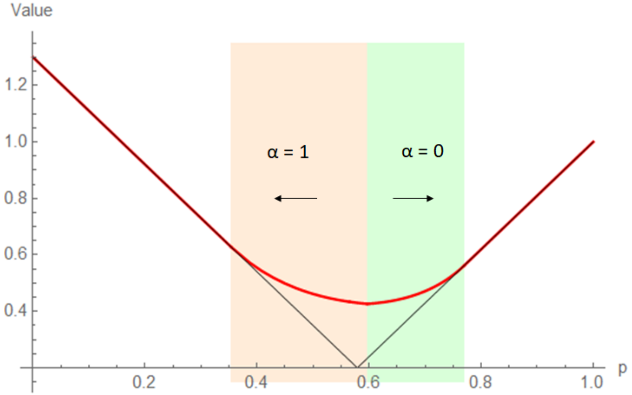
P's Individually Optimal Learning with low r

$r < \underline{r}(\Delta) \Rightarrow$ "contradictory" and "confirmatory" learning



P's Individually Optimal Learning with intermediate r

$$\overline{r}(\Delta) > r \geq \underline{r}(\Delta) \Rightarrow \text{ “contradictory” learning}$$



Individually optimal learning regions

$$\big|\text{————————}\underline{p}^{*P} \leftarrow\leftarrow\leftarrow\leftarrow \check{p}^{P} \rightarrow\rightarrow\rightarrow\rightarrow \bar{p}^{*P}\text{————————}\big|$$

$$\underset{p=0}{\big|}\ \underbrace{\text{—————}}_{\text{decision } b}\ \underline{p}^{*E}\ \underbrace{\leftarrow\leftarrow\leftarrow\leftarrow}_{\alpha=1}\ \check{p}^{E}\ \underbrace{\rightarrow\rightarrow\rightarrow\rightarrow}_{\alpha=0}\ \bar{p}^{*E}\ \underbrace{\text{——————————}}_{\text{decision } a}\ \underset{p=1}{\big|}$$

Figure: Individually Optimal Learning

- P & E agree on the "regime" for optimal learning
- P & E may disagree on the source of information or whether to learn at a given p
- E's optimal learning region is shifted to the right relative to P

Low r

Complete Delegation Payoff

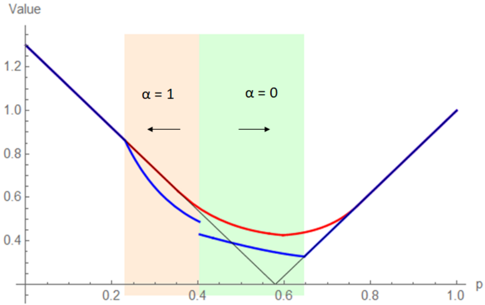
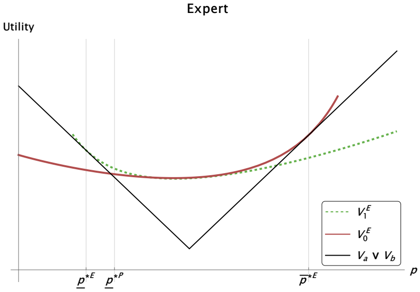


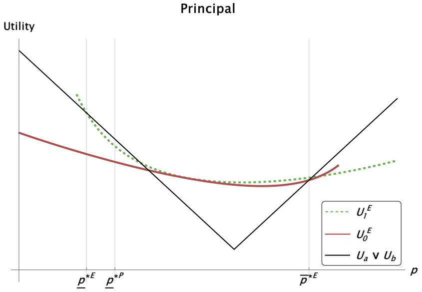
Figure: Complete delegation payoff

- Under complete delegation, P's value function is discontinuous
- For some low p, complete delegation is worse than immediate decision

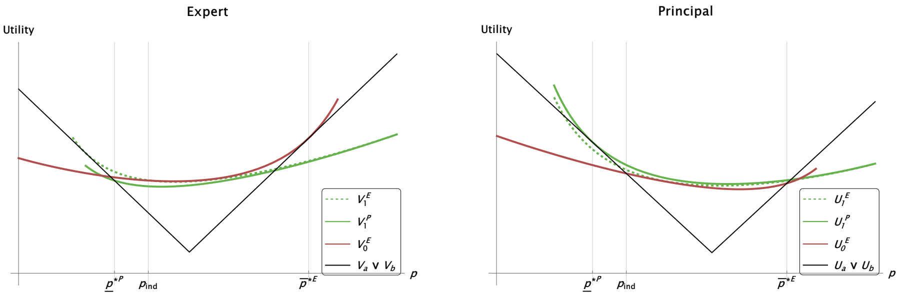
Low r

Flexible v Complete Delegation: Low Δ



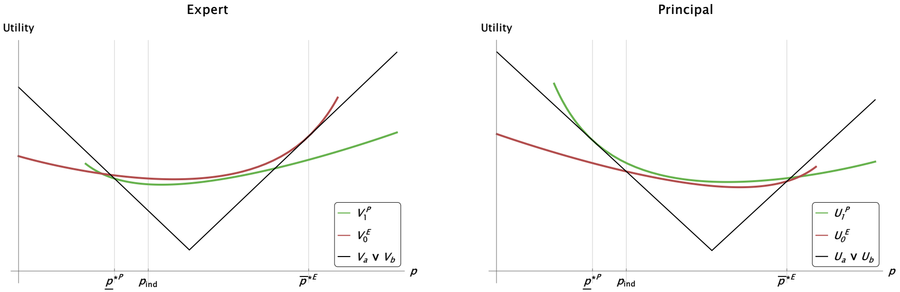


Flexible v Complete Delegation: Low Δ



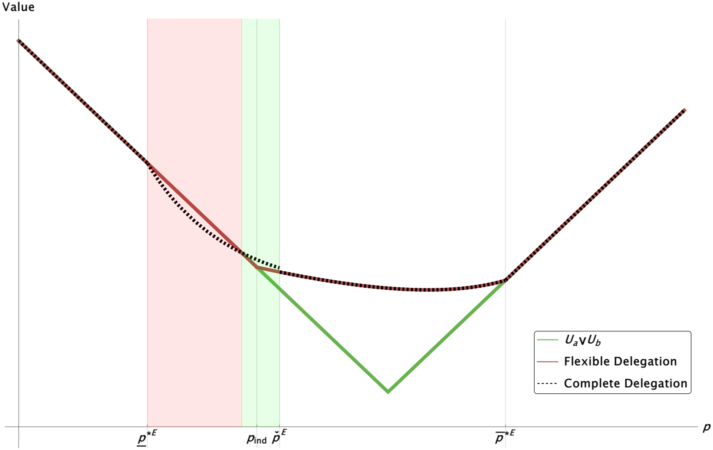
- P can stop learning at low p when immediate decision is optimal
- P would like to, but cannot continue learning at higher $p \Rightarrow$ Interventions are one-sided
- Benefit: No unnecessary delay
- Cost: E's strategic response $\Rightarrow$ For small Δ cost exceeds benefit for some p

Flexible v Complete Delegation: Low Δ



- P stops A-learning at $\underline{p}^{*P}$ and allows B-learning from p_{ind}. For $p \in (\underline{p}^{*P}, p_{ind})$ E sets $\alpha_t \in (0,1)$ so as to keep P indifferent.

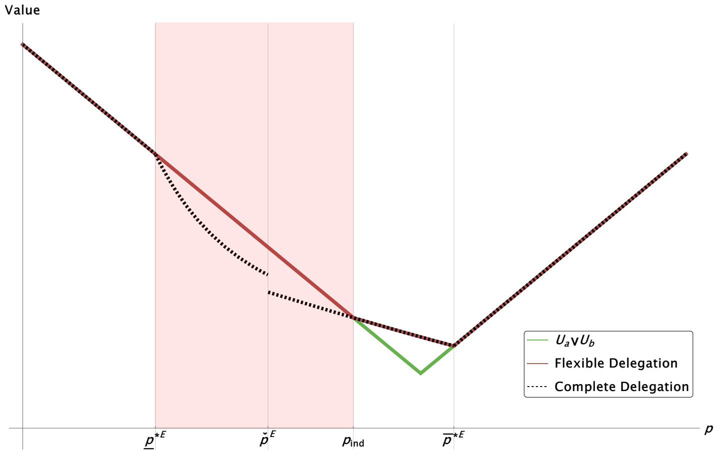
Complete Delegation SOMETIMES optimal



Gain: P's value never below the value from immediate decision

Loss: E's learning even more sub-optimal for P

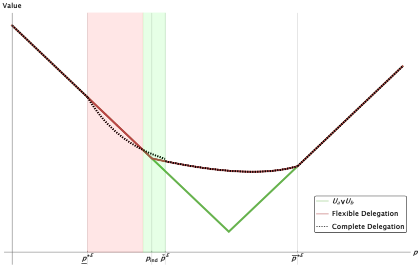
Complete Delegation NEVER optimal



<u>Gain</u>: P's value never below the value from immediate action

<u>Loss</u>: E's strategic response does not affect P

Flexible v Complete Delegation



Low Δ $\Rightarrow$ E A-learns when P's value of learning +ve

High Δ $\Rightarrow$ E A-learns when P's value of learning $-$ve

Concluding Remarks

- The need to retain ultimate authority over decision making arises naturally in the environment to keep the biased expert in check.
- Retaining some decision making authority could be damaging for the principal if and only if the expert's preferences are close the principal's.

Thank You!

HJB Formulation

$$
\begin{aligned}
rV(p) \geqslant \max_{\alpha} \Big\{ & \mathbb{I}\{Y(p,\alpha)=\emptyset\}\left(p\alpha\left(v_a^A - V(p)\right)\right) + (1-p)(1-\alpha)\left(v_b^B - V(p)\right) \\
& - (2\alpha-1)\,p\,(1-p)\,V'(p) + \mathbb{I}\{Y(p,\alpha)=a\}\,V_a(p) \\
& + \mathbb{I}\{Y(p,\alpha)=b\}\,V_b(p) \Big\}, \quad p \in (0,1)
\end{aligned}
$$

subject to the boundary condition

$$
V\left(p^{\sigma}(\tau)\right) = V_a\left(p^{\sigma}(\tau)\right) \vee V_b\left(p^{\sigma}(\tau)\right)
$$

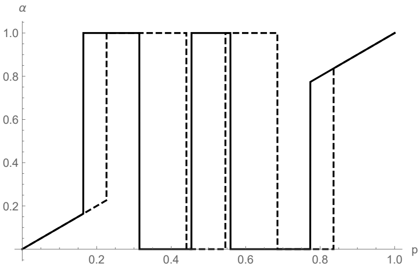
and the constraint

$$
U(p) \geq U_a(p) \vee U_b(p)
$$

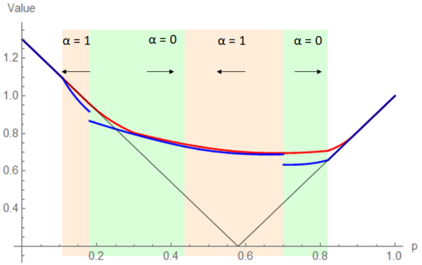
Back

FB v Complete Delegation: Low r

$r < \underline{r}(\Delta) \Rightarrow$ "contradictory" and "confirmatory" learning by both P & E



Optimal Learning



Value Functions

Back