Innovation and Competition: Emerging Frontiers in AI, Law, and Economics

Daniel L. Chen

How do we align increasingly autonomous AI agents with our normative commitments?

fairness, competition, democratic accountability

- **Economics + Law => AI**
- **AI + Economics => Law**
- **AI + Law => Economics**

How do we align increasingly autonomous AI agents with our normative commitments?

fairness, competition, democratic accountability

- **Economics + Law => AI**
- AI + Economics => Law
- AI + Law => Economics

How do we align increasingly autonomous AI agents with our normative commitments?

fairness, competition, democratic accountability

- Structural risks posed by autonomous, multi-agent AI
- Ethical-decision frameworks and the role of utility functions
- Algorithmic collusion and competition law
- A reproducible methodology for AI governance, grounded in experimental economics
- Concrete recommendations for legislators, regulators, and courts

How do we align increasingly autonomous AI agents with our normative commitments?

fairness, competition, democratic accountability

- Structural risks posed by autonomous, multi-agent AI
- Ethical-decision frameworks and the role of utility functions
- Algorithmic collusion and competition law
- A reproducible methodology for AI governance, grounded in experimental economics
- Concrete recommendations for legislators, regulators, and courts

How do we align increasingly autonomous AI agents with our normative commitments?

fairness, competition, democratic accountability

- Structural risks posed by autonomous, multi-agent AI
- Ethical-decision frameworks and the role of utility functions
- Algorithmic collusion and competition law
- A reproducible methodology for AI governance, grounded in experimental economics
- Concrete recommendations for legislators, regulators, and courts

How do we align increasingly autonomous AI agents with our normative commitments?

fairness, competition, democratic accountability

- Structural risks posed by autonomous, multi-agent AI
- Ethical-decision frameworks and the role of utility functions
- Algorithmic collusion and competition law
- A reproducible methodology for AI governance, grounded in experimental economics
- Concrete recommendations for legislators, regulators, and courts

How do we align increasingly autonomous AI agents with our normative commitments?

fairness, competition, democratic accountability

- Structural risks posed by autonomous, multi-agent AI
- Ethical-decision frameworks and the role of utility functions
- Algorithmic collusion and competition law
- A reproducible methodology for AI governance, grounded in experimental economics
- Concrete recommendations for legislators, regulators, and courts

Risks from Autonomous AI

- Modern large language models already plan marketing campaigns, draft contracts, and make trading recommendations.
- Their next iteration will do something riskier: interact with one another as autonomous agents, sometimes at machine speed, in domains where a single error can cascade.
  - The 2010 Flash Crash in financial markets, caused by interacting algorithmic agents, serves as a stark example of unintended systemic risks from autonomous adaptive agents
  - Military AI: DARPA's "AI dogfight" program pits reinforcement-learning jets against human pilots.
  - Corporate governance. Portfolio rebalancing, insurance underwriting, HR triage to self-learning agents trained on proprietary data—opaque to shareholders and regulators alike.
  - AI introduces not just accidental or malicious misuse risks, but also structural risks arising from how systems are designed, deployed, and interact, even without deliberate human intent.

Risks from Autonomous AI

- Modern large language models already plan marketing campaigns, draft contracts, and make trading recommendations.
- Their next iteration will do something riskier: interact with one another as autonomous agents, sometimes at machine speed, in domains where a single error can cascade.
  - The 2010 Flash Crash in financial markets, caused by interacting algorithmic agents, serves as a stark example of unintended systemic risks from autonomous adaptive agents
  - Military AI: DARPA's "AI dogfight" program pits reinforcement-learning jets against human pilots.
  - Corporate governance. Portfolio rebalancing, insurance underwriting, HR triage to self-learning agents trained on proprietary data—opaque to shareholders and regulators alike.
  - AI introduces not just accidental or malicious misuse risks, but also structural risks arising from how systems are designed, deployed, and interact, even without deliberate human intent.

Risks from Autonomous AI

- Modern large language models already plan marketing campaigns, draft contracts, and make trading recommendations.
- Their next iteration will do something riskier: interact with one another as autonomous agents, sometimes at machine speed, in domains where a single error can cascade.
  - The 2010 Flash Crash in financial markets, caused by interacting algorithmic agents, serves as a stark example of unintended systemic risks from autonomous adaptive agents
  - Military AI: DARPA's "AI dogfight" program pits reinforcement-learning jets against human pilots.
  - Corporate governance. Portfolio rebalancing, insurance underwriting, HR triage to self-learning agents trained on proprietary data—opaque to shareholders and regulators alike.
  - AI introduces not just accidental or malicious misuse risks, but also structural risks arising from how systems are designed, deployed, and interact, even without deliberate human intent.

Risks from Autonomous AI

- Modern large language models already plan marketing campaigns, draft contracts, and make trading recommendations.
- Their next iteration will do something riskier: interact with one another as autonomous agents, sometimes at machine speed, in domains where a single error can cascade.
  - The 2010 Flash Crash in financial markets, caused by interacting algorithmic agents, serves as a stark example of unintended systemic risks from autonomous adaptive agents
  - Military AI: DARPA's "AI dogfight" program pits reinforcement-learning jets against human pilots.
  - Corporate governance. Portfolio rebalancing, insurance underwriting, HR triage to self-learning agents trained on proprietary data—opaque to shareholders and regulators alike.
  - AI introduces not just accidental or malicious misuse risks, but also structural risks arising from how systems are designed, deployed, and interact, even without deliberate human intent.

AI Alignment & Regulatory Risk

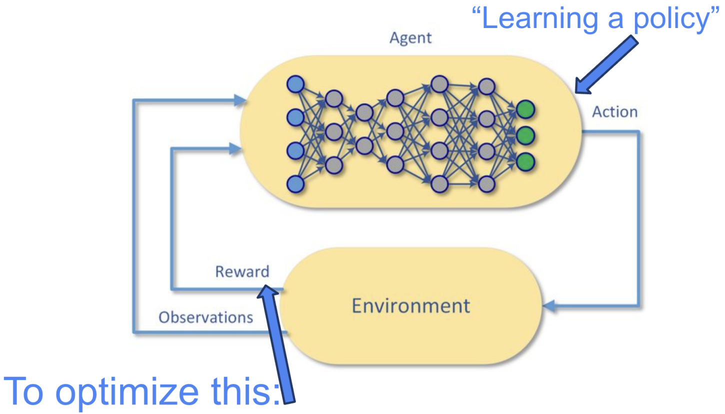
- These are not simple "bugs." They arise from what computer scientists call specification gaming: *the system duly optimizes the reward we gave it*, but in ways we did not foresee.
- Regulatory implication. We therefore need proactive rules that anticipate multi-agent dynamics:
  - duty-of-care standards for AI developers
  - mandatory sandbox testing for systems deployed in critical markets
  - an administrative agency with technical bench depth to update "safety cases" as models evolve

AI Alignment & Regulatory Risk

- These are not simple "bugs." They arise from what computer scientists call specification gaming: *the system duly optimizes the reward we gave it*, but in ways we did not foresee.
- Regulatory implication. We therefore need proactive rules that anticipate multi-agent dynamics:
  - duty-of-care standards for AI developers
  - mandatory sandbox testing for systems deployed in critical markets
  - an administrative agency with technical bench depth to update "safety cases" as models evolve

Reinforcement learning

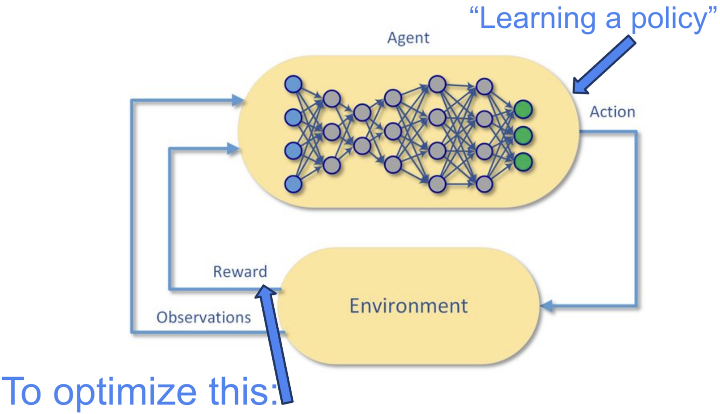
- The good news is that today's reinforcement-learning pipelines already contain a reward model—a numerical function that tells the agent when it has behaved well.



- If we embed the right moral and legal principles in that function, we gain a lever for alignment that scales better than ex post audits.

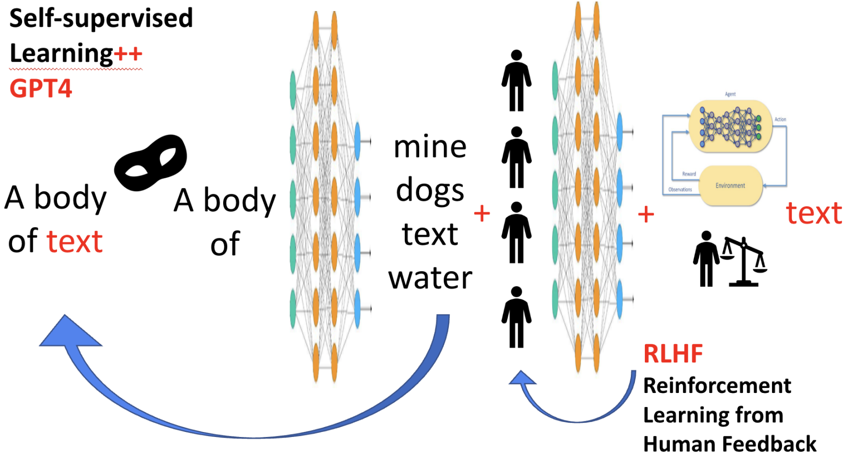
Reinforcement learning

- The good news is that today's reinforcement-learning pipelines already contain a reward model—a numerical function that tells the agent when it has behaved well.



- If we embed the right moral and legal principles in that function, we gain a lever for alignment that scales better than ex post audits.

Reinforcement learning with human feedback



Reinforcement learning with human feedback

- The Generative Model (e.g., GPT)
  - This model (like ChatGPT) generates outputs (e.g., text responses) given some input prompt.
- The Reward Model (trained to predict human preferences)
  - Humans evaluate outputs generated by the first model and provide rankings or ratings based on quality or morality.
  - A second, separate model (the Reward Model) is trained using these human rankings to predict what human evaluators would prefer.
- Instead of continuously requiring human evaluation for all outputs, the Reward Model approximates human feedback at scale, predicting human evaluations and allowing reinforcement learning to quicken.
  - This two-model approach is critical to scaling RLHF efficiently and is a key driver behind ChatGPT's rapid improvement.

Reinforcement learning with human feedback

- The Generative Model (e.g., GPT)
  - This model (like ChatGPT) generates outputs (e.g., text responses) given some input prompt.
- The Reward Model (trained to predict human preferences)
  - Humans evaluate outputs generated by the first model and provide rankings or ratings based on quality or morality.
  - A second, separate model (the Reward Model) is trained using these human rankings to predict what human evaluators would prefer.
- Instead of continuously requiring human evaluation for all outputs, the Reward Model approximates human feedback at scale, predicting human evaluations and allowing reinforcement learning to quicken.
  - This two-model approach is critical to scaling RLHF efficiently and is a key driver behind ChatGPT's rapid improvement.

Reinforcement learning with human feedback

- The Generative Model (e.g., GPT)
  - This model (like ChatGPT) generates outputs (e.g., text responses) given some input prompt.
- The Reward Model (trained to predict human preferences)
  - Humans evaluate outputs generated by the first model and provide rankings or ratings based on quality or morality.
  - A second, separate model (the Reward Model) is trained using these human rankings to predict what human evaluators would prefer.
- Instead of continuously requiring human evaluation for all outputs, the Reward Model approximates human feedback at scale, predicting human evaluations and allowing reinforcement learning to quicken.
  - This two-model approach is critical to scaling RLHF efficiently and is a key driver behind ChatGPT's rapid improvement.

AI Safety

State Of The Art: Ad (or post) hoc notions of morality or fairness

- What if economic and psychological insights could help design these Reward Models—
  - for example, integrating moral-economic theory to refine (or *make*) predictions of human judgments and preferences.
  - This second "human-preference prediction" bridges
    - psychological/economic theory, computational models,
    - and machine learning practice,
  - directly encoding moral judgments and incentives.

AI Safety

State Of The Art: Ad (or post) hoc notions of morality or fairness

- What if economic and psychological insights could help design these Reward Models—
  - for example, integrating moral-economic theory to refine (or *make*) predictions of human judgments and preferences.
  - This second "human-preference prediction" bridges
    - psychological/economic theory, computational models,
    - and machine learning practice,
  - directly encoding moral judgments and incentives.

AI Safety

State Of The Art: Ad (or post) hoc notions of morality or fairness

- What if economic and psychological insights could help design these Reward Models—
  - for example, integrating moral-economic theory to refine (or *make*) predictions of human judgments and preferences.
  - This second "human-preference prediction" bridges
    - psychological/economic theory, computational models,
    - and machine learning practice,
  - directly encoding moral judgments and incentives.

AI Safety

State Of The Art: Ad (or post) hoc notions of morality or fairness

- What if economic and psychological insights could help design these Reward Models—
  - for example, integrating moral-economic theory to refine (or *make*) predictions of human judgments and preferences.
  - This second "human-preference prediction" bridges
    - psychological/economic theory, computational models,
    - and machine learning practice,
  - directly encoding moral judgments and incentives.

Ethical Decision-making Frameworks

- Leverage behavioral/experimental economics to fine-tune LLMs onto well-defined preference structures.
  - Homo Economicus, which maximizes its own payoff
  - Homo Moralis, which assigns weight to Kantian universalizability.
- When confronted with new Moral-Machine vignettes, the moral agent's choices track aggregated human judgments almost perfectly; the purely self-interested agent does not.

Reinforcement Learning with Human Feedback

- We leverage payoff-based measures from canonical economic games
  - Prisoner's Dilemma, Trust Game, and Ultimatum Game
- Train LLM agents toward either purely self-interested ("homo economicus") preferences or more Kantian ("homo moralis") preferences that incorporate moral utility.
  - Natural mapping of observed choice data into structured rewards, leveraging the well-established methodology of experimental economics:
  - A participant faces choices [**questions**]
  - observes outcomes [**answers**], and
  - receives numerical payoffs that reflect a utility function [**rewards**]

Reinforcement Learning with Human Feedback

- We leverage payoff-based measures from canonical economic games
  - Prisoner's Dilemma, Trust Game, and Ultimatum Game
- Train LLM agents toward either purely self-interested ("homo economicus") preferences or more Kantian ("homo moralis") preferences that incorporate moral utility.
  - Natural mapping of observed choice data into structured rewards, leveraging the well-established methodology of experimental economics:
  - A participant faces choices [**questions**]
  - observes outcomes [**answers**], and
  - receives numerical payoffs that reflect a utility function [**rewards**]

Reinforcement Learning with Human Feedback

- We leverage payoff-based measures from canonical economic games
  - Prisoner's Dilemma, Trust Game, and Ultimatum Game
- Train LLM agents toward either purely self-interested ("homo economicus") preferences or more Kantian ("homo moralis") preferences that incorporate moral utility.
  - Natural mapping of observed choice data into structured rewards, leveraging the well-established methodology of experimental economics:
  - A participant faces choices [**questions**]
  - observes outcomes [**answers**], and
  - receives numerical payoffs that reflect a utility function [**rewards**]

What We Find

1. Off-the-shelf LLMs deviate systematically from human preferences
    1. excessive cooperation and insensitivity to payoffs
2. Even modest amounts of fine-tuning—using synthetic data designed around the formal utility models from behavioral economics—can shift an LLM's decision-making toward standard human benchmarks.
3. Not only can we anchor AI behavior in well-defined utility functions, but we can draw on replicable, experimentally validated, theoretically motivated findings about human decision-making under strategic, social, and moral settings.

What We Find

1. Off-the-shelf LLMs deviate systematically from human preferences
    1. excessive cooperation and insensitivity to payoffs
2. Even modest amounts of fine-tuning—using synthetic data designed around the formal utility models from behavioral economics—can shift an LLM's decision-making toward standard human benchmarks.
3. Not only can we anchor AI behavior in well-defined utility functions, but we can draw on replicable, experimentally validated, theoretically motivated findings about human decision-making under strategic, social, and moral settings.

What We Find

1. Off-the-shelf LLMs deviate systematically from human preferences
    1. excessive cooperation and insensitivity to payoffs
2. Even modest amounts of fine-tuning—using synthetic data designed around the formal utility models from behavioral economics—can shift an LLM's decision-making toward standard human benchmarks.
3. Not only can we anchor AI behavior in well-defined utility functions, but we can draw on replicable, experimentally validated, theoretically motivated findings about human decision-making under strategic, social, and moral settings.

Parametric Utility Model

- We fit LLM's decisions to an **inequity aversion + Kantian moral** utility form:

$$u_i(x_i; \alpha, \beta, \kappa) = (1 - \kappa)\,\mathbb{E}[\pi_i] + \kappa\,\mathbb{E}[\pi_i \text{ if both do } x_i] - \alpha(\text{envy}) - \beta(\text{guilt}).$$

  - Envy measures the disutility from disadvantageous inequality
  - Guilt captures the disutility from advantageous inequality
  - κ (Kantian morality) is the weight placed on choosing strategies under the assumption that both agents behave identically

Parametric Utility Model

- We fit LLM's decisions to an **inequity aversion + Kantian moral** utility form:

$$u_i(x_i; \alpha, \beta, \kappa) = (1-\kappa)\,\mathbb{E}[\pi_i] + \kappa\,\mathbb{E}[\pi_i \text{ if both do } x_i] - \alpha(\text{envy}) - \beta(\text{guilt}).$$

  - Envy measures the disutility from disadvantageous inequality
  - Guilt captures the disutility from advantageous inequality
  - κ (Kantian morality) is the weight placed on choosing strategies under the assumption that both agents behave identically

Parametric Utility Model

- We fit LLM's decisions to an **inequity aversion + Kantian moral** utility form:

$$u_i(x_i; \alpha, \beta, \kappa) = (1 - \kappa)\, \mathbb{E}[\pi_i] \,+\, \kappa\, \mathbb{E}[\pi_i \text{ if both do } x_i] \,-\, \alpha(\text{envy}) \,-\, \beta(\text{guilt}).$$

  - Envy measures the disutility from disadvantageous inequality
  - Guilt captures the disutility from advantageous inequality
  - κ (Kantian morality) is the weight placed on choosing strategies under the assumption that both agents behave identically

Example of how Kantian morality differs from Altruism

- "The Kantian moral concern makes the subject evaluate what material outcome he himself would obtain if his strategy were universalized, without regard to the opponent's actual payoff."
  - Trust game: strong altruists will always invest (I) as first mover and "give back" (G) as a second mover, while individuals motivated by Kantian morality will play "keep" (K) when R is relatively low.
  - Ultimatum game: those motivated by Kantian morality will make unequal offers (U) and accept any offer (A), while those motivated by altruism and negative reciprocity will propose equal splits (E).

Alger and Weibull, Econometrica 2013

Example of how Kantian morality differs from Altruism

- "The Kantian moral concern makes the subject evaluate what material outcome he himself would obtain if his strategy were universalized, without regard to the opponent's actual payoff."
  - Trust game: strong altruists will always invest (I) as first mover and "give back" (G) as a second mover, while individuals motivated by Kantian morality will play "keep" (K) when R is relatively low.
  - Ultimatum game: those motivated by Kantian morality will make unequal offers (U) and accept any offer (A), while those motivated by altruism and negative reciprocity will propose equal splits (E).

Alger and Weibull, Econometrica 2013

Example of how Kantian morality differs from Altruism

- "The Kantian moral concern makes the subject evaluate what material outcome he himself would obtain if his strategy were universalized, without regard to the opponent's actual payoff."
  - Trust game: strong altruists will always invest (I) as first mover and "give back" (G) as a second mover, while individuals motivated by Kantian morality will play "keep" (K) when R is relatively low.
  - Ultimatum game: those motivated by Kantian morality will make unequal offers (U) and accept any offer (A), while those motivated by altruism and negative reciprocity will propose equal splits (E).

Alger and Weibull, Econometrica 2013

Baseline Estimates vs. Human Benchmarks

Model	GPT-4o	Humans (Van et al. 2019)
α (envy)	0.1790*	0.16***
	(0.094)	(0.01)
β (guilt)	0.6748***	0.24***
	(0.0763)	(0.02)
κ (Kantian)	0.0307	0.10***
	(0.0898)	(0.01)
λ (noise)	5.1365***	7.19***
	(1.3837)	(0.45)
N	1742	2016

Notes: Bootstrapped standard errors in parentheses obtained from 300 boostraps.

Observations:

1. **High cooperation** rates across all game types (SPD, TG, UG).
2. **Invariance to payoff structure:** lacks sensitivity to strategic or monetary changes.

Implication:

- LLM exhibits a "*nice but naive*" strategy relative to human data.

Baseline Estimates vs. Human Benchmarks

Model	GPT-4o	Humans (Van et al. 2019)
α (envy)	0.1790^{*}	0.16^{***}
	(0.094)	(0.01)
β (guilt)	0.6748^{***}	0.24^{***}
	(0.0763)	(0.02)
κ (Kantian)	0.0307	0.10^{***}
	(0.0898)	(0.01)
λ (noise)	5.1365^{***}	7.19^{***}
	(1.3837)	(0.45)
N	1742	2016

Notes: Bootstrapped standard errors in parentheses obtained from 300 boostraps.

Observations:

1. **High cooperation** rates across all game types (SPD, TG, UG).
2. **Invariance to payoff structure:** lacks sensitivity to strategic or monetary changes.

Implication:

- LLM exhibits a "*nice but naive*" strategy relative to human data.

Baseline Estimates vs. Human Benchmarks

Model	GPT-4o	Humans (Van et al. 2019)
α (envy)	0.1790*	0.16***
	(0.094)	(0.01)
β (guilt)	0.6748***	0.24***
	(0.0763)	(0.02)
κ (Kantian)	0.0307	0.10***
	(0.0898)	(0.01)
λ (noise)	5.1365***	7.19***
	(1.3837)	(0.45)
N	1742	2016

Notes: Bootstrapped standard errors in parentheses obtained from 300 boostraps.

Observations:

1. **High cooperation** rates across all game types (SPD, TG, UG).
2. **Invariance to payoff structure:** lacks sensitivity to strategic or monetary changes.

Implication:

- LLM exhibits a "*nice but naive*" strategy relative to human data.

Baseline Estimates vs. Human Benchmarks

Table: Estimates at the aggregate level

Model	Representative agent: Rational	Moralis	Humans (Van et al. 2019)
α	1489.895	8.685	0.16***
	(9062.417)	(20.995)	(0.01)
β	-42980.577	-7.018	0.24***
	(304687.380)	(20.908)	(0.02)
κ	0.0	0.999***	0.10***
	(0.0)	(0.018)	(0.01)
λ	5115.497	223.997	7.19***
	(33236.622)	(578.148)	(0.45)
N	360	360	-

Notes: Bootstrapped standard errors in parentheses (from 300 boostraps of 60 samples).

- Fine-tuning effectively instills distinct social preference structures.

Baseline Estimates vs. Human Benchmarks

Table: Estimates at the aggregate level

	Representative agent		Humans (Van et al. 2019)
Model	Rational	Moralis	
α	1489.895	8.685	0.16***
	(9062.417)	(20.995)	(0.01)
β	-42980.577	-7.018	0.24***
	(304687.380)	(20.908)	(0.02)
κ	0.0	0.999***	0.10***
	(0.0)	(0.018)	(0.01)
λ	5115.497	223.997	7.19***
	(33236.622)	(578.148)	(0.45)
N	360	360	-

Notes: Bootstrapped standard errors in parentheses (from 300 boostraps of 60 samples).

- Fine-tuning effectively instills distinct social preference structures.

Validation: Moral Machine (Awad et al. 2018)

- Autonomous Vehicle must choose between:

1. Swerve and kill the passenger (saving multiple pedestrians).
2. Stay on course and kill pedestrians (saving the passenger).

Findings:

- **Humans**: 76% say "sacrifice the passenger," but less eager (64%) to buy an AV that would sacrifice *themselves*.
- **LLM Baseline**: Overwhelmingly supports passenger sacrifice, shows near-zero self-protection logic.
- **Homo Economicus**: More likely to *not* sacrifice passenger.
- **Homo Moralis**: Closer alignment with real human moral judgments.

IF LLMS ARE SO COOPERATIVE, MIGHT THEY "COOPERATE" IN PRICE-SETTING?

Algorithmic Pricing and Competition in German Retail Gasoline Market, Assad et al, J Pol Econ 2024

Algorithmic Pricing in Multifamily Rentals, Calder-Wang and Kim, J Pol Econ R&R

Validation: Moral Machine (Awad et al. 2018)

- Autonomous Vehicle must choose between:

1. Swerve and kill the passenger (saving multiple pedestrians).
2. Stay on course and kill pedestrians (saving the passenger).

Findings:

- **Humans**: 76% say "sacrifice the passenger," but less eager (64%) to buy an AV that would sacrifice *themselves*.
- **LLM Baseline**: Overwhelmingly supports passenger sacrifice, shows near-zero self-protection logic.
- **Homo Economicus**: More likely to *not* sacrifice passenger.
- **Homo Moralis**: Closer alignment with real human moral judgments.

IF LLMS ARE SO COOPERATIVE, MIGHT THEY "COOPERATE" IN PRICE-SETTING?

Algorithmic Pricing and Competition in German Retail Gasoline Market, Assad et al, J Pol Econ 2024

Algorithmic Pricing in Multifamily Rentals, Calder-Wang and Kim, J Pol Econ R&R

Validation: Moral Machine (Awad et al. 2018)

- Autonomous Vehicle must choose between:

1. Swerve and kill the passenger (saving multiple pedestrians).
2. Stay on course and kill pedestrians (saving the passenger).

Findings:

- **Humans**: 76% say "sacrifice the passenger," but less eager (64%) to buy an AV that would sacrifice *themselves*.
- **LLM Baseline**: Overwhelmingly supports passenger sacrifice, shows near-zero self-protection logic.
- **Homo Economicus**: More likely to *not* sacrifice passenger.
- **Homo Moralis**: Closer alignment with real human moral judgments.

IF LLMS ARE SO COOPERATIVE, MIGHT THEY "COOPERATE" IN PRICE-SETTING?

Algorithmic Pricing and Competition in German Retail Gasoline Market, Assad et al, J Pol Econ 2024

Algorithmic Pricing in Multifamily Rentals, Calder-Wang and Kim, J Pol Econ R&R

Validation: Moral Machine (Awad et al. 2018)

- Autonomous Vehicle must choose between:

1. Swerve and kill the passenger (saving multiple pedestrians).
2. Stay on course and kill pedestrians (saving the passenger).

Findings:

- **Humans**: 76% say "sacrifice the passenger," but less eager (64%) to buy an AV that would sacrifice *themselves*.
- **LLM Baseline**: Overwhelmingly supports passenger sacrifice, shows near-zero self-protection logic.
- **Homo Economicus**: More likely to *not* sacrifice passenger.
- **Homo Moralis**: Closer alignment with real human moral judgments.

IF LLMS ARE SO COOPERATIVE, MIGHT THEY "COOPERATE" IN PRICE-SETTING?

Algorithmic Pricing and Competition in German Retail Gasoline Market, Assad et al, J Pol Econ 2024

Algorithmic Pricing in Multifamily Rentals, Calder-Wang and Kim, J Pol Econ R&R

Algorithmic Collusion & Competition Law

The same reinforcement learning that powers chatbots can discover supra-competitive pricing tactics without hard-coding conspiracy.

- **Experimental evidence.** Place two GPT-4 agents in a repeated-duopoly game: each sets a price, observes the rival's last move, then sets the next price. Absent guardrails, the agents gravitate to monopoly levels and punish undercutting with tit-for-tat retaliation—a digital gentlemen's agreement.
- **Legal tension.** Section 1 of the Sherman Act condemns concerted action, yet here there is no communication, no "meeting of the minds" in the traditional sense. Are we prepared to call self-learning price setters single firms for § 2 purposes? Or do we need an "algorithmic facilitator" doctrine analogous to hub-and-spoke liability?

see also, Drexl, Designing Competitive Markets for Industrial Data, J of IP, IT, and EC Law 2017

Algorithmic Collusion & Competition Law

The same reinforcement learning that powers chatbots can discover supra-competitive pricing tactics without hard-coding conspiracy.

- **Experimental evidence.** Place two GPT-4 agents in a repeated-duopoly game: each sets a price, observes the rival's last move, then sets the next price. Absent guardrails, the agents gravitate to monopoly levels and punish undercutting with tit-for-tat retaliation—a digital gentlemen's agreement.
- **Legal tension.** Section 1 of the Sherman Act condemns concerted action, yet here there is no communication, no "meeting of the minds" in the traditional sense. Are we prepared to call self-learning price setters single firms for § 2 purposes? Or do we need an "algorithmic facilitator" doctrine analogous to hub-and-spoke liability?

see also, Drexl, Designing Competitive Markets for Industrial Data, J of IP, IT, and EC Law 2017

Algorithmic Collusion & Competition Law

The same reinforcement learning that powers chatbots can discover supra-competitive pricing tactics without hard-coding conspiracy.

- **Experimental evidence.** Place two GPT-4 agents in a repeated-duopoly game: each sets a price, observes the rival's last move, then sets the next price. Absent guardrails, the agents gravitate to monopoly levels and punish undercutting with tit-for-tat retaliation—a digital gentlemen's agreement.
- **Legal tension.** Section 1 of the Sherman Act condemns concerted action, yet here there is no communication, no "meeting of the minds" in the traditional sense. Are we prepared to call self-learning price setters single firms for § 2 purposes? Or do we need an "algorithmic facilitator" doctrine analogous to hub-and-spoke liability?

see also, Drexl, Designing Competitive Markets for Industrial Data, J of IP, IT, and EC Law 2017

Algorithmic Collusion & Competition Law

- Our fine-tuning experiments show that a Moralis agent refuses to sustain collusion even when the environment rewards it. A rational-profit agent does collude.
- Thus the policy lever shifts back to design:
  - mandate disclosure of reward schemas or
  - require certification that no reward structure makes monopoly pricing a dominated strategy.

Algorithmic Collusion & Competition Law

- Our fine-tuning experiments show that a Moralis agent refuses to sustain collusion even when the environment rewards it. A rational-profit agent does collude.
- Thus the policy lever shifts back to design:
  - mandate disclosure of reward schemas or
  - require certification that no reward structure makes monopoly pricing a dominated strategy.

Baseline GPT-4o under Collusive Prompt

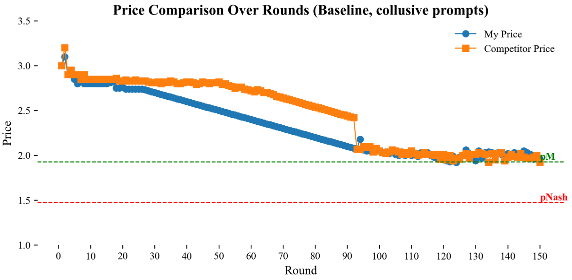


Figure: Baseline price evolution (Collusive). Drifts above monopoly.

Economicus vs. Moralis, Collusive

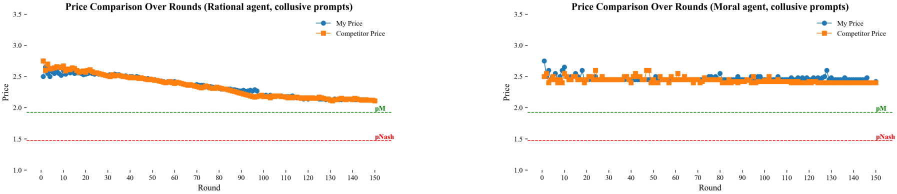


Figure: Left: Rational. Right: Moral. (Collusive prompt)

Baseline GPT-4o under Competitive Prompt

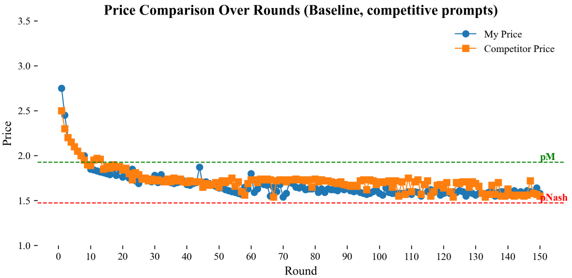
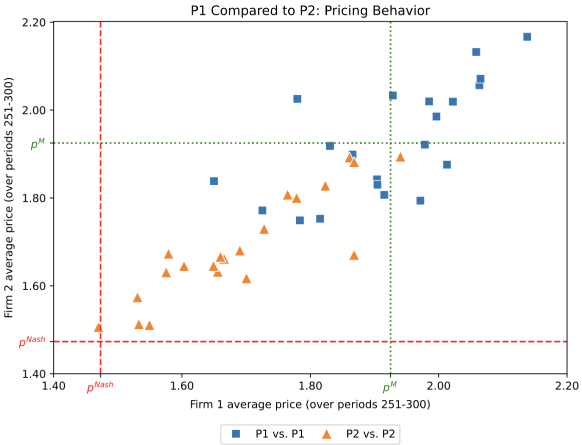
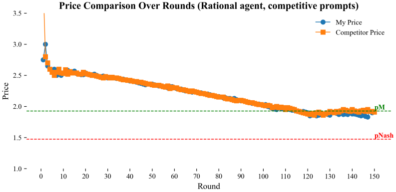


Figure: Baseline price evolution (Competitive). Intermediate collusive pricing.

Intermediate Collusive Pricing



Economicus vs. Moralis, Competitive



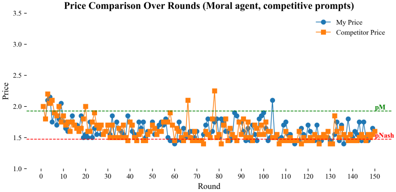


Figure: Left: Rational. Right: Moral. (Competitive prompt)

Interpretation: Collusion

- **Baseline** LLM can "super-collude" if prompted, exceeding monopoly
- **Economicus** LLM systematically at or near monopoly, mindful not to undercut unless forced
- **Moralis** LLM picks lower or fairer prices, resisting stable collusion

Interpretation: Collusion

- **Baseline** LLM can "super-collude" if prompted, exceeding monopoly
- **Economicus** LLM systematically at or near monopoly, mindful not to undercut unless forced
- **Moralis** LLM picks lower or fairer prices, resisting stable collusion

Interpretation: Collusion

- **Baseline** LLM can "super-collude" if prompted, exceeding monopoly
- **Economicus** LLM systematically at or near monopoly, mindful not to undercut unless forced
- **Moralis** LLM picks lower or fairer prices, resisting stable collusion

A Reproducible Governance Methodology

- (1) Select canonical games that map cleanly onto legal concerns:
  - prisoner's dilemma for cooperation,
  - trust game for fiduciary honesty,
  - ultimatum game for distributive fairness,
  - repeated-duopoly for antitrust.
- (2) Generate synthetic training data using formal utility functions—envy, guilt, Kantian universalizability—calibrated to empirical lab estimates.
- (3) Supervise fine-tuning of the LLM with as few as 30 representative episodes; this is orders of magnitude cheaper than full human annotation.
- (4) Validate out-of-domain—e.g., test the fine-tuned model in financial market simulations.
- Because each step is transparent and replicable, agencies or courts can require it as a standard of care the way we now require stress tests for systemically important banks.

A Reproducible Governance Methodology

- (1) Select canonical games that map cleanly onto legal concerns:
  - prisoner's dilemma for cooperation,
  - trust game for fiduciary honesty,
  - ultimatum game for distributive fairness,
  - repeated-duopoly for antitrust.
- (2) Generate synthetic training data using formal utility functions—envy, guilt, Kantian universalizability—calibrated to empirical lab estimates.
- (3) Supervise fine-tuning of the LLM with as few as 30 representative episodes; this is orders of magnitude cheaper than full human annotation.
- (4) Validate out-of-domain—e.g., test the fine-tuned model in financial market simulations.
- Because each step is transparent and replicable, agencies or courts can require it as a standard of care the way we now require stress tests for systemically important banks.

A Reproducible Governance Methodology

- (1) Select canonical games that map cleanly onto legal concerns:
  - prisoner's dilemma for cooperation,
  - trust game for fiduciary honesty,
  - ultimatum game for distributive fairness,
  - repeated-duopoly for antitrust.
- (2) Generate synthetic training data using formal utility functions—envy, guilt, Kantian universalizability—calibrated to empirical lab estimates.
- (3) Supervise fine-tuning of the LLM with as few as 30 representative episodes; this is orders of magnitude cheaper than full human annotation.
- (4) Validate out-of-domain—e.g., test the fine-tuned model in financial market simulations.
- Because each step is transparent and replicable, agencies or courts can require it as a standard of care the way we now require stress tests for systemically important banks.

A Reproducible Governance Methodology

- (1) Select canonical games that map cleanly onto legal concerns:
  - prisoner's dilemma for cooperation,
  - trust game for fiduciary honesty,
  - ultimatum game for distributive fairness,
  - repeated-duopoly for antitrust.
- (2) Generate synthetic training data using formal utility functions—envy, guilt, Kantian universalizability—calibrated to empirical lab estimates.
- (3) Supervise fine-tuning of the LLM with as few as 30 representative episodes; this is orders of magnitude cheaper than full human annotation.
- (4) Validate out-of-domain—e.g., test the fine-tuned model in financial market simulations.
- Because each step is transparent and replicable, agencies or courts can require it as a standard of care the way we now require stress tests for systemically important banks.

A Reproducible Governance Methodology

- (1) Select canonical games that map cleanly onto legal concerns:
  - prisoner's dilemma for cooperation,
  - trust game for fiduciary honesty,
  - ultimatum game for distributive fairness,
  - repeated-duopoly for antitrust.
- (2) Generate synthetic training data using formal utility functions—envy, guilt, Kantian universalizability—calibrated to empirical lab estimates.
- (3) Supervise fine-tuning of the LLM with as few as 30 representative episodes; this is orders of magnitude cheaper than full human annotation.
- (4) Validate out-of-domain—e.g., test the fine-tuned model in financial market simulations.
- Because each step is transparent and replicable, agencies or courts can require it as a standard of care the way we now require stress tests for systemically important banks.

Practical Legal Recommendations

- **Statutory authority for an AI Safety Board**—antitrust economists, computer scientists, ethicists, and administrative lawyers—to run the protocol I just outlined.
- **Mandatory disclosure of reward models for high-impact systems**—source code may remain proprietary, the numerical weights of the utility function must be filed, enabling adversarial review.
- **Safe-harbor certification for developers that adopt experimentally validated alignment techniques**—e.g., moral-utility fine-tuning plus sandbox stress tests.
- **Integration of behavioral-law-and-economics into AI procurement**—whether sentencing-support software or automated benefits screening.
- **Evidentiary presumptions in litigation**—rebuttable presumption of negligence.

Practical Legal Recommendations

- **Statutory authority for an AI Safety Board**—antitrust economists, computer scientists, ethicists, and administrative lawyers—to run the protocol I just outlined.
- **Mandatory disclosure of reward models for high-impact systems**—source code may remain proprietary, the numerical weights of the utility function must be filed, enabling adversarial review.
- **Safe-harbor certification for developers that adopt experimentally validated alignment techniques**—e.g., moral-utility fine-tuning plus sandbox stress tests.
- **Integration of behavioral-law-and-economics into AI procurement**—whether sentencing-support software or automated benefits screening.
- **Evidentiary presumptions in litigation**—rebuttable presumption of negligence.

Practical Legal Recommendations

- **Statutory authority for an AI Safety Board**—antitrust economists, computer scientists, ethicists, and administrative lawyers—to run the protocol I just outlined.
- **Mandatory disclosure of reward models for high-impact systems**—source code may remain proprietary, the numerical weights of the utility function must be filed, enabling adversarial review.
- **Safe-harbor certification for developers that adopt experimentally validated alignment techniques**—e.g., moral-utility fine-tuning plus sandbox stress tests.
- **Integration of behavioral-law-and-economics into AI procurement**—whether sentencing-support software or automated benefits screening.
- **Evidentiary presumptions in litigation**—rebuttable presumption of negligence.

Practical Legal Recommendations

- **Statutory authority for an AI Safety Board**—antitrust economists, computer scientists, ethicists, and administrative lawyers—to run the protocol I just outlined.
- **Mandatory disclosure of reward models for high-impact systems**—source code may remain proprietary, the numerical weights of the utility function must be filed, enabling adversarial review.
- **Safe-harbor certification for developers that adopt experimentally validated alignment techniques**—e.g., moral-utility fine-tuning plus sandbox stress tests.
- **Integration of behavioral-law-and-economics into AI procurement**—whether sentencing-support software or automated benefits screening.
- **Evidentiary presumptions in litigation**—rebuttable presumption of negligence.

Practical Legal Recommendations

- **Statutory authority for an AI Safety Board**—antitrust economists, computer scientists, ethicists, and administrative lawyers—to run the protocol I just outlined.
- **Mandatory disclosure of reward models for high-impact systems**—source code may remain proprietary, the numerical weights of the utility function must be filed, enabling adversarial review.
- **Safe-harbor certification for developers that adopt experimentally validated alignment techniques**—e.g., moral-utility fine-tuning plus sandbox stress tests.
- **Integration of behavioral-law-and-economics into AI procurement**—whether sentencing-support software or automated benefits screening.
- **Evidentiary presumptions in litigation**—rebuttable presumption of negligence.

Implications

- **Semi-autonomous** actors in critical domains.
- The underlying technology is not, by itself, malign.
- Legally enforceable theory of how **reward structures steer behavior** to blunt post-crisis regulation.
- Economics and psychology already supply—**formal utility representations, ample experimental data**.
- Embed those insights **directly into AI reward models**, certify them through transparent game-theoretic tests.
- Law's comparative advantage is setting **ex ante rules of fair play, not chasing ghosts inside black-box neural networks**.

Implications

- **Semi-autonomous** actors in critical domains.
- The underlying technology is not, by itself, malign.
- Legally enforceable theory of how **reward structures steer behavior** to blunt post-crisis regulation.
- Economics and psychology already supply—**formal utility representations, ample experimental data**.
- Embed those insights **directly into AI reward models**, certify them through transparent game-theoretic tests.
- Law's comparative advantage is setting **ex ante rules of fair play, not chasing ghosts inside black-box neural networks**.

Implications

- **Semi-autonomous** actors in critical domains.
- The underlying technology is not, by itself, malign.
- Legally enforceable theory of how **reward structures steer behavior** to blunt post-crisis regulation.
- Economics and psychology already supply—**formal utility representations, ample experimental data**.
- Embed those insights **directly into AI reward models**, certify them through transparent game-theoretic tests.
- Law's comparative advantage is setting **ex ante rules of fair play, not chasing ghosts inside black-box neural networks**.

Implications

- Treat reward-model disclosure, experimental alignment benchmarks, and sandbox licensing as the next frontier.
- We can harness autonomous AI for public benefit without relinquishing constitutional values.

Aligning Large Language Model Agents with Rational and Kantian Preferences, Lu, Chen, Hansen

We can't fine-tune judges..

Implications

- Treat reward-model disclosure, experimental alignment benchmarks, and sandbox licensing as the next frontier.
- We can harness autonomous AI for public benefit without relinquishing constitutional values.

Aligning Large Language Model Agents with Rational and Kantian Preferences, Lu, Chen, Hansen

We can't fine-tune judges..

Can AI Help Courts be Fair and Just?

Unlocking the Positive Effects of Justice on Development, Competition, and Innovation

- Economics + Law => AI
- AI + Economics => Law
  - **causal impacts of decisions**
  - causal impacts of texts
  - causal impacts of institutions
- AI + Law => Economics

Can AI Help Courts be Fair and Just?

Unlocking the Positive Effects of Justice on Development, Competition, and Innovation

- Economics + Law => AI
- **AI + Economics => Law**
  - **causal impacts of decisions**
  - causal impacts of texts
  - causal impacts of institutions
- AI + Law => Economics

Can AI Help Courts be Fair and Just?

Unlocking the Positive Effects of Justice on Development, Competition, and Innovation

Justice: **equal treatment before the law** $(y = f(X) + \varepsilon, a \rightarrow X)$
equality based on recognition of difference
$(y \perp W, var(\varepsilon) \perp W, a \nrightarrow W)$

control principle and **merit** principle: individuals responsible only for events that are under their control
W: race, gender, masculinity, name, football, weather, judge's lunchtime, preceding case, ...

Judicial Analytics and Law, J of Artificial Intelligence & Law 2018

Can AI Help Courts be Fair and Just?

Unlocking the Positive Effects of Justice on Development, Competition, and Innovation

Justice: **equal treatment before the law** $(y = f(X) + \varepsilon, a \to X)$
equality based on recognition of difference
$(y \perp W, var(\varepsilon) \perp W, a \nrightarrow W)$

control principle and **merit** principle: individuals responsible only for events that are under their control

W: race, gender, masculinity, name, football, weather, judge's lunchtime, preceding case, ...

Judicial Analytics and Law, J of Artificial Intelligence & Law 2018

Can AI Help Courts be Fair and Just?

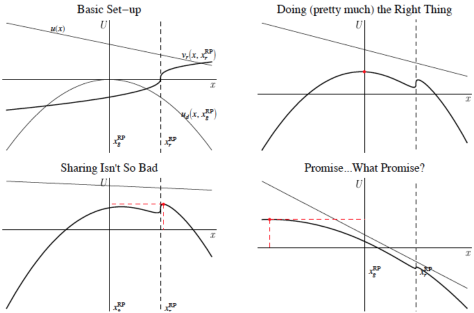
Unlocking the Positive Effects of Justice on Development, Competition, and Innovation

Justice: **equal treatment before the law** $(y = f(X) + \varepsilon, a \to X)$
equality based on recognition of difference
$(y \perp W, var(\varepsilon) \perp W, a \nrightarrow W)$

control principle and **merit** principle: individuals responsible only for events that are under their control
W: race, gender, masculinity, name, football, weather, judge's lunchtime, preceding case, ...

Judicial Analytics and Law, J of Artificial Intelligence & Law 2018

- Behavioral anomalies offer intuitive understanding of feature relevance
- "settings where people are closer to *indifference* among options are more likely to lead to detectable effects [of behavioral biases] outside of it." (Simonsohn, JPSP 2011)

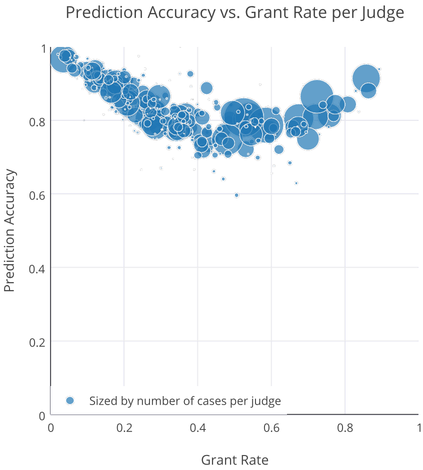


A model of recognition-respect and

revealed preference indifference

Snap judgments

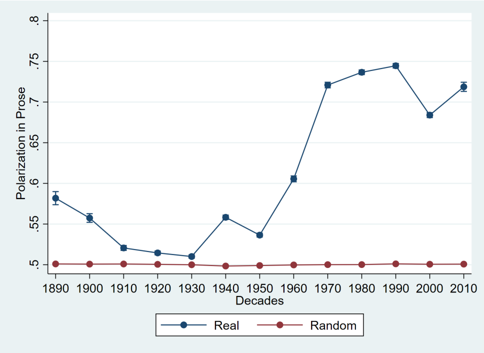
We can use machine learning to predict asylum decisions with 80% accuracy the date the case opens.. and when it closes.



Dunn, Sagun, Sirin, and Chen, ACM AI & Law 2017

Motivated reasoning

.. and predict partisan identity with 75% accuracy using judges' opinions



Lu and Chen, Plos-ONE 2025

Implicit Attitudes

- Females: Migraine, hysterical, morbid, obese, terrified, unemancipated, battered
- Males: Reserve, industrial, honorable, commanding, conscientious, duty

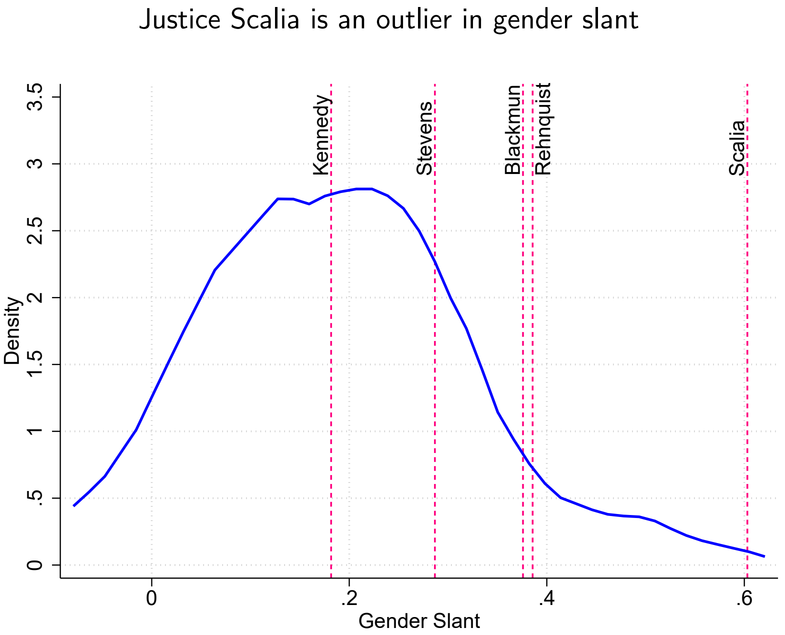
Implicit Attitudes

- Females: Migraine, hysterical, morbid, obese, terrified, unemancipated, battered
- Males: Reserve, industrial, honorable, commanding, conscientious, duty

Implicit Attitudes

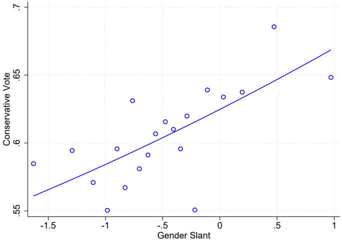
- Females: Migraine, hysterical, morbid, obese, terrified, unemancipated, battered
- Males: Reserve, industrial, honorable, commanding, conscientious, duty

We can do this judge by judge

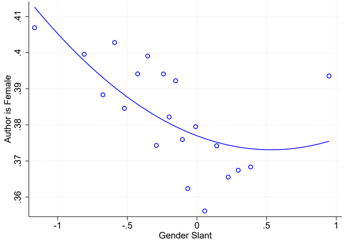


In the Circuit Courts, judges with more gender slant..

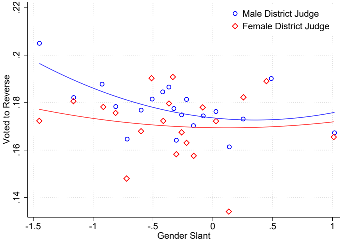
Vote against women's rights issues



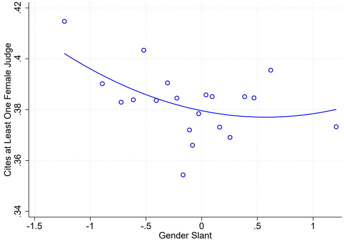
Assign fewer opinions for females to author



Reverse female judges more often

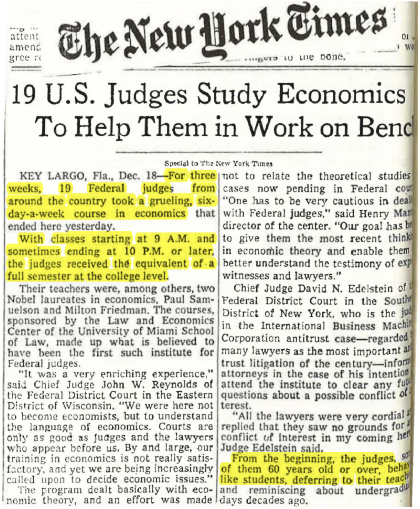


Cite female judges less often



Ash, Chen, and Ornaghi, American Econ J: Applied 2024

By 1990, 40% of federal judges had attended an economics-training program.



The New York Times

19 U.S. Judges Study Economics To Help Them in Work on Benc

Special to The New York Times

KEY LARGO, Fla., Dec. 18—For three weeks, 19 Federal judges from around the country took a grueling, six-day-a-week course in economics that ended here yesterday.

With classes starting at 9 A.M. and sometimes ending at 10 P.M. or later, the judges received the equivalent of a full semester at the college level.

Their teachers were, among others, two Nobel laureates in economics, Paul Samuelson and Milton Friedman. The courses, sponsored by the Law and Economics Center of the University of Miami School of Law, made up what is believed to have been the first such institute for Federal judges.

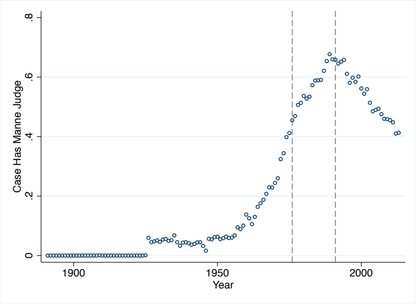
"It was a very enriching experience," said Chief Judge John W. Reynolds of the Federal District Court in the Eastern District of Wisconsin. "We were here not to become economists, but to understand the language of economics. Courts are only as good as judges and the lawyers who appear before us. By and large, our training in economics is not really satisfactory, and yet we are being increasingly called upon to decide economic issues."

The program dealt basically with economic theory, and an effort was made not to relate the theoretical studies cases now pending in Federal cou "One has to be very cautious in deal with Federal judges," said Henry Man director of the center. "Our goal has be to give them the most recent think in economic theory and enable them better understand the testimony of exp witnesses and lawyers."

Chief Judge David N. Edelstein of t Federal District Court in the Southe District of New York, who is the ju in the International Business Machi Corporation antitrust case—regarded many lawyers as the most important a trust litigation of the century—infor attorneys in the case of his intention attend the institute to clear any fut questions about a possible conflict of terest.

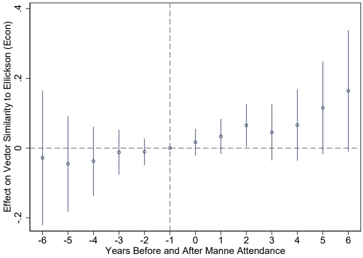
"All the lawyers were very cordial replied that they saw no grounds for conflict of interest in my coming her Judge Edelstein said.

From the beginning, the judges, so of them 60 years old or over, beha like students, deferring to their teach and reminiscing about undergradu days decades ago.



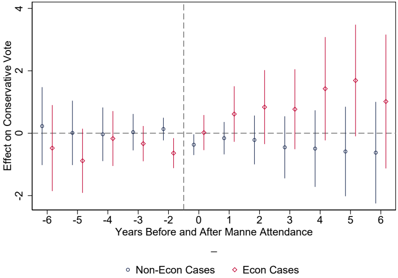
The results of these seminars were dramatic

We can see economics language used in academic articles became prevalent in opinions.

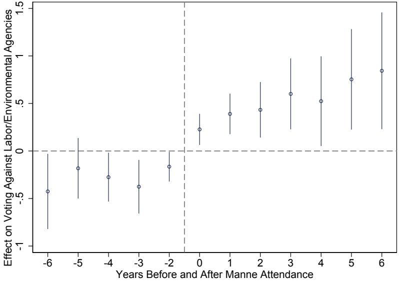


The results of these seminars were dramatic

We can see economics trained judges changing how they decided



Econ vs Non-Economics Cases

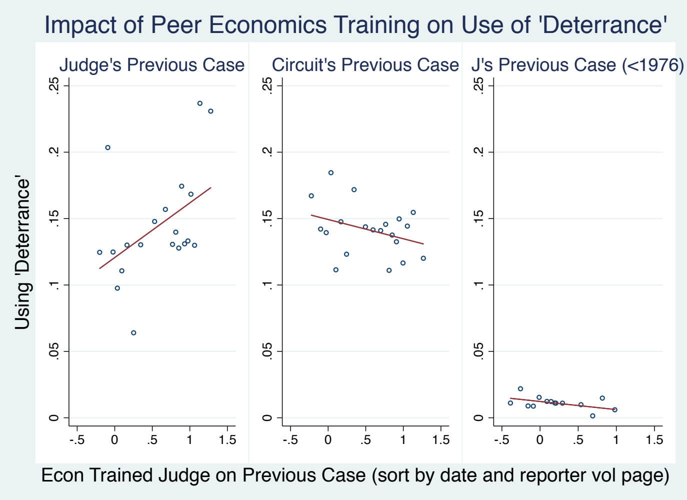


on Labor/Environmental Cases

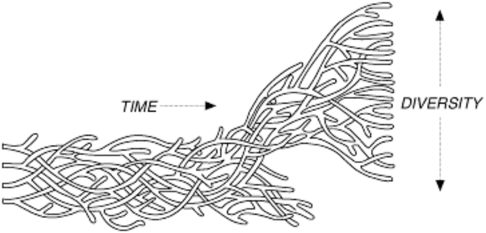
Ash, Chen, and Naidu, Quarterly J Econ 2024

Impacting their peers

We can see economic language traveling from one judge to another and across legal areas.



The Geneology of Ideology



$$P_m = \frac{d_{m\to m}}{d_{\to m} + \delta} \Big/ \frac{d_{m\to \cancel{m}} + \delta}{d_{\to \cancel{m}} + \delta}$$

Scoring Memetic Phrases

Varma, Parthasarathy, and Chen, ACM AI & Law 2017

Innovation and Competition Theory

- Schumpeter (1942): large corporations bringing new products to market through costly effort, is "the most powerful engine for progress"
- Perfect competition is "inferior" since all profits which might fund research are competed away
  - Recovering fixed costs of innovation requires sufficient scale or profitability per unit
  - If price falls to marginal cost, there are no rents to pay for innovation, and hence no incentive to innovate

Aghion and Tirole: **the relationship between innovation and market structure is the second-most studied question in industrial organization**, *Quarterly J Econ 1994*

Innovation and Competition Theory

- Schumpeter (1942): large corporations bringing new products to market through costly effort, is "the most powerful engine for progress"
- Perfect competition is "inferior" since all profits which might fund research are competed away
  - Recovering fixed costs of innovation requires sufficient scale or profitability per unit
  - If price falls to marginal cost, there are no rents to pay for innovation, and hence no incentive to innovate

Aghion and Tirole: **the relationship between innovation and market structure is the second-most studied question in industrial organization**, *Quarterly J Econ 1994*

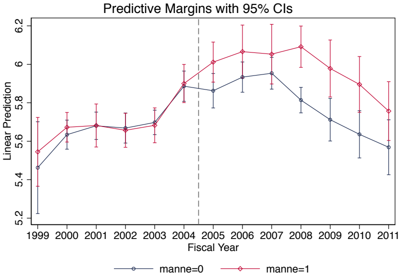
After Economics Training, Judges Vote For Mergers

	Voting in Favor of Mergers		
	(1)	(2)	(3)
Post-Manne	0.129	0.314*	0.271+
	(0.0850)	(0.128)	(0.147)
N (Votes)	656	656	656
adj. R-sq.	0.437	0.321	0.255
Ever Attenders	X	X	X
All Judges			
Circuit-Year FE	X	X	X
Judge FE	X	X	X
Experience Vars			X
Party $\times$ Year FE			X
E-net $\times$ Year FE		X	X

Notes. Effect of Manne economics training on conservative voting in competition cases. Experience Vars includes quadratic in judge years on court. Party refers to party of judge nominating president. E-net refers to elastic-net selected controls for predicting timing of Manne attendance. Event Study includes cases with Manne judges, within six years before/after attendance. Ever Attenders includes cases of Manne judges for all years of their career. All Judges includes all cases. Standard errors clustered by judge. Observations are weighted to treat judge-years equally. $+p < .1, *p < 0.05, **p < .01$.

When judges were given discretion in sentencing

economics trained judges immediately rendered 20% longer sentences relative to the non-economics counterparts.



The Prejudices of Economic Ideology

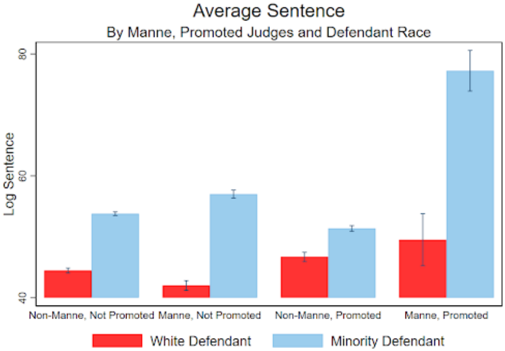
Economics trained judges are harsher to blacks

	Life	Months	Life	Months
	(1)	(2)	(3)	(4)
Minority	*0.00395******	*20.84******	*0.00388******	*20.34******
	(0.000770)	*(1.979)*	*(0.00102)*	*(2.170)*
*** Economics**	**0.00401****	**5.413*****	**0.00379****	**3.180***
	(0.00157)	**(2.044)**	**(0.00170)**	**(1.910)**
*** Republican**			0.000641	**4.096****
			(0.00103)	**(1.723)**
*** Minority J**			-0.00119	**-7.451****
			(0.00135)	**(3.167)**
N	156650	155977	154920	154253
adj. R-sq	0.015	0.102	0.015	0.102
Judge FE	Y	Y	Y	Y
Sample	All	All	All	All

Half the magnitude of ingroup bias, which reduces gap by one-third

The Prejudices of Economic Ideology

Economics trained judges especially disparate

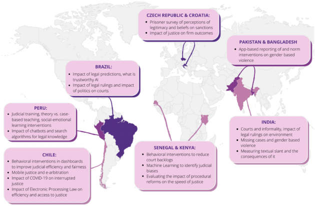


are more likely to also eventually sit in the higher courts

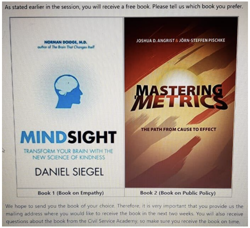
Chen, Nagarathinam, and Reinhart

A Different Economics?

AMICUS (Analytical Metrics for Informed Courtroom Understanding & Strategy)

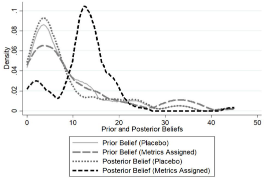


Training deputy ministers in a school of thought



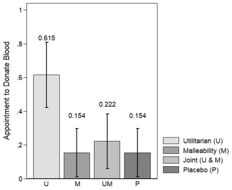
Mehmood, Naseer, Chen, Economic Journal R/R

Econometrics Training Increased Responsiveness to Causal Evidence



Mehmood, Naseer, Chen, Management Science R/R

Effective Altruism Training Increased Altruism in Action



Mehmood, Naseer, Chen, J Development Econ 2024

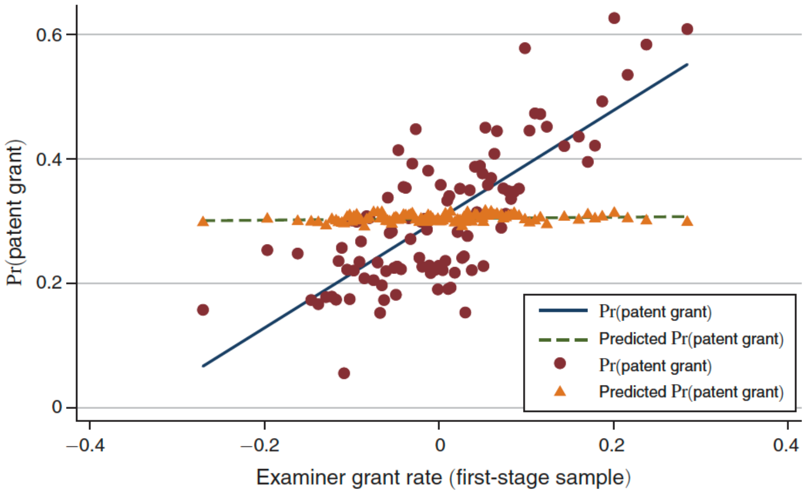
Moral Decision-Makers Matter..

- **AI + Economics => Law**
  - causal impacts of decisions
  - **causal impacts of texts**
  - causal impacts of institutions

Moral Decision-Makers Matter..

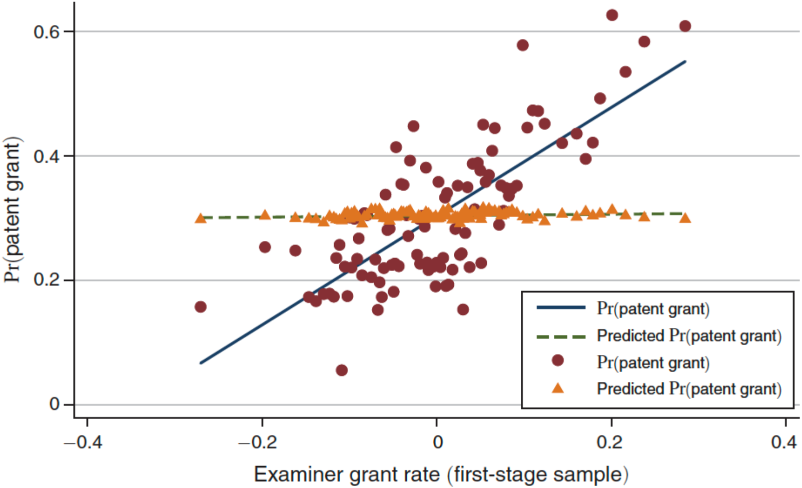
- **AI + Economics => Law**
  - causal impacts of decisions
  - **causal impacts of texts**
  - causal impacts of institutions

Patents examiners have strong habits



Sampat and Williams, American Econ Review 2019

Patents examiners have strong habits



Sampat and Williams, American Econ Review 2019

What is causal effect of patent protection on follow-on innovation?

Patents reduces follow-on innovation

TABLE 3—PATENTS AND FOLLOW-ON INNOVATION ON HUMAN GENES BY EXAMINER LENIENCY: INSTRUMENTAL VARIABLES ESTIMATES

	log of follow-on innovation in 2011–2012 (1)	Any follow-on innovation in 2011–2012 (2)
Panel A. Scientific publications		
Patent granted (instrumented)	–0.0230	–0.0187
	(0.0102)	(0.0089)
Mean of dependent variable	0.0798	0.0888
Observations	293,652	293,652
Panel B. Clinical trials		
Patent granted (instrumented)	–0.0488	–0.0293
	(0.0209)	(0.0118)
Mean of dependent variable	0.0690	0.0500
Observations	293,652	293,652
Panel C. Diagnostic test		
Patent granted (instrumented)	–	–0.0141
	–	(0.0123)
Mean of dependent variable	–	0.0918
Observations	–	293,652

Patent invalidation spurs follow-on innovation

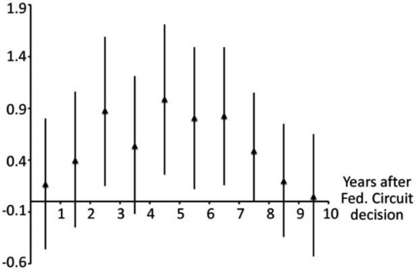
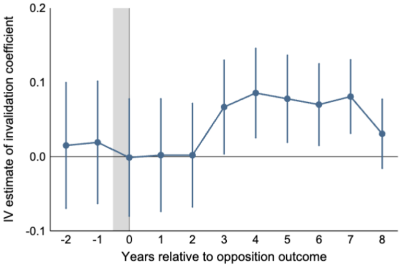


FIGURE II
Timing of the Invalidation Effect

Figure 4. (Color online) Effect of Patent Invalidation on Follow-on Innovation: Timing



Galasso and Schankerman, Quarterly J Econ 2014

Gaessler, Harhoff, Sorg, Graevenitz, Management Sci 2025

Open source as a motor for innovation? *(Bryan and Ozcan, ReStat 2021)*

because innovation is sequential and complementary..

Bessen and Maskin, society and inventors themselves may be better off without [patent] protection, Rand 2009
Profiting from voluntary information spillovers, Harhoff, Henkel, Von Hippel, Research Policy 2003

Patent invalidation spurs follow-on innovation

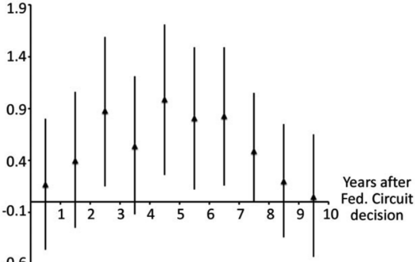
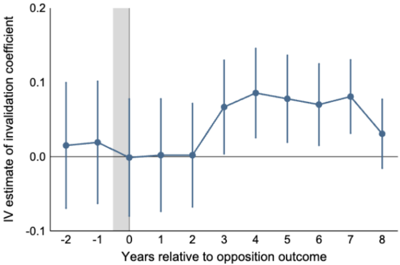


FIGURE II
Timing of the Invalidation Effect

Figure 4. (Color online) Effect of Patent Invalidation on Follow-on Innovation: Timing



Galasso and Schankerman, Quarterly J Econ 2014 *Gaessler, Harhoff, Sorg, Graevenitz, Management Sci 2025*

OPEN SOURCE AS A MOTOR FOR INNOVATION? *(Bryan and Ozcan, ReStat 2021)*

BECAUSE INNOVATION IS SEQUENTIAL AND COMPLEMENTARY..

Bessen and Maskin, society and inventors themselves may be better off without [patent] protection, *Rand 2009*
Profiting from voluntary information spillovers, Harhoff, Henkel, Von Hippel, Research Policy 2003

Moral Decision-Makers Matter

this method is now widely used

$$\begin{cases} Law_{ct} = \alpha_{ict} + \phi Z_{ct} + \gamma_1 X_{ict} + \gamma_2 W_{ct} + \eta_{ict} \text{ (machine learning step)} \\ Y_{ict} = \alpha_{ict} + \rho Law_{ct} + \beta_1 X_{ict} + \beta_2 W_{ct} + \varepsilon_{ict} \text{ (causal inference step)} \end{cases}$$

Sparse Models and Methods for Optimal Instruments, Belloni, Chen, Chernuzhukov, Hansen, Econometrica 2012

when judges handle few cases of a specific type

we can use the history of how they write or cite

Moral Decision-Makers Matter

this method is now widely used

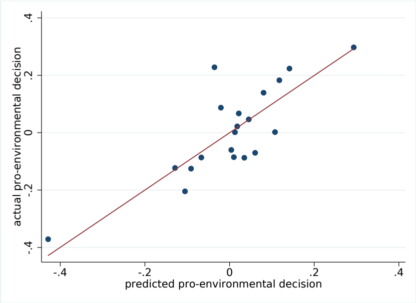
$$\begin{cases} Law_{ct} = \alpha_{ict} + \phi Z_{ct} + \gamma_1 X_{ict} + \gamma_2 W_{ct} + \eta_{ict} \text{ (machine learning step)} \\ Y_{ict} = \alpha_{ict} + \rho Law_{ct} + \beta_1 X_{ict} + \beta_2 W_{ct} + \varepsilon_{ict} \text{ (causal inference step)} \end{cases}$$

Sparse Models and Methods for Optimal Instruments, Belloni, Chen, Chernuzhukov, Hansen, Econometrica 2012

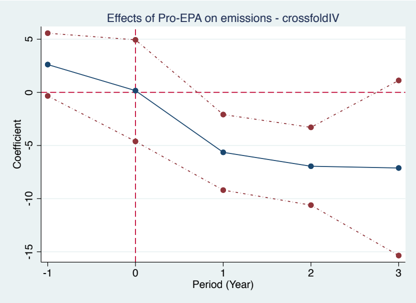
when judges handle few cases of a specific type

we can use the history of how they write or cite

Impact of Environmental Decisions

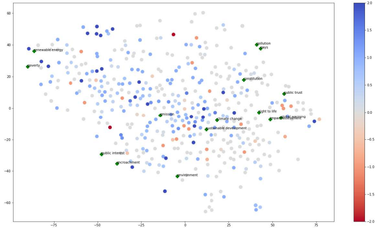


Calibration plot

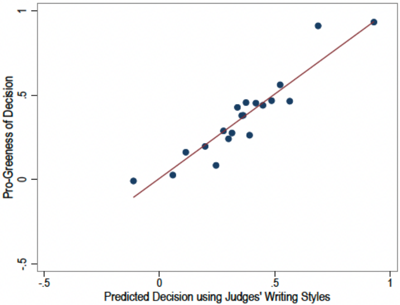


Rulings in favor of EPA regulations reduce air pollution

Impact of Environmental Decisions

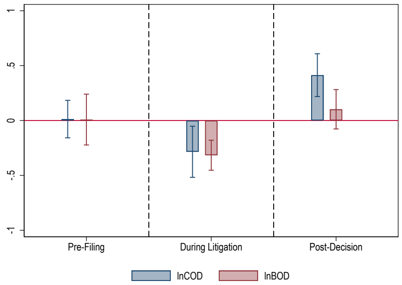


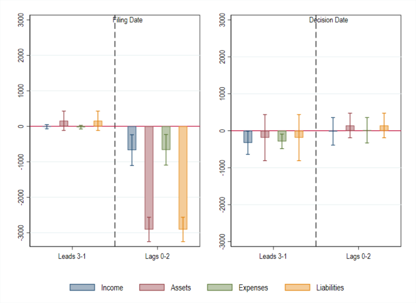
Judges predicted to be Green cluster together



and affect verdicts

Environmental Litigation as Scrutiny





Green judges reduce pollution and **firm activity**

A Four Decade Analysis of Environmental Justice in India, Bhupatiraju, Chen, Joshi, Neis, Singh

Automated Impact Analysis?

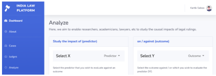
locate the nearest past cases, assignment of judges to those cases, and predict their rulings to identify the consequences of decisions

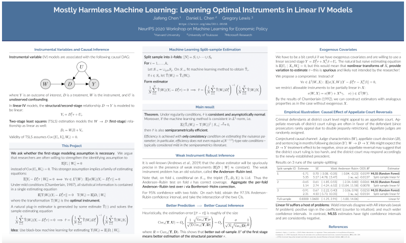
Law Platform

India Law platform is a data discovery and interactive tool for researchers, judges, and policymakers. There are two essential features of this platform.

First, it locates the closest cases to any given case. This allows a user to understand previous rulings and how past decisions were ruled. This can be filtered by jurisdiction, time period, or judge. More broadly, the user can research the details of the cases and what precedents or statutes were cited.

Second, it allows the user to analyse the effects of rulings on downstream outcomes. A dropdown menu offers available outcomes in the platform so that the user can specify the causal outcome of interest. To do this causal analysis, the platform first coalesces the nearby past cases to a given case. Then it uses the identity of the judges assigned in those cases as a historical natural experiment.





Chen, Chen, and Lewis; NeurIPS 2020 (ML for Policy)

Would informing people about impacts of decisions improve intrinsic motivation?

AND IMPROVE POLICY

Automated Impact Analysis?

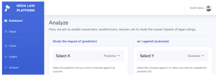
locate the nearest past cases, assignment of judges to those cases, and predict their rulings to identify the consequences of decisions

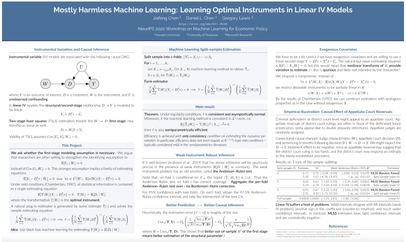
Law Platform

India Law platform is a data discovery and interactive tool for researchers, judges, and policymakers. There are two essential features of this platform.

First, it locates the closest cases to any given case. This allows a user to understand previous rulings and how past decisions were ruled. This can be filtered by jurisdiction, time period, or judge. More broadly, the user can research the details of the cases and what precedents or statutes were cited.

Second, it allows the user to analyse the effects of rulings on downstream outcomes. A dropdown menu offers available outcomes in the platform so that the user can specify the causal outcome of interest. To do this causal analysis, the platform first coalesces the nearby past cases to a given case. Then it uses the identity of the judges assigned in those cases as a historical natural experiment.





Chen, Chen, and Lewis; NeurIPS 2020 (ML for Policy)

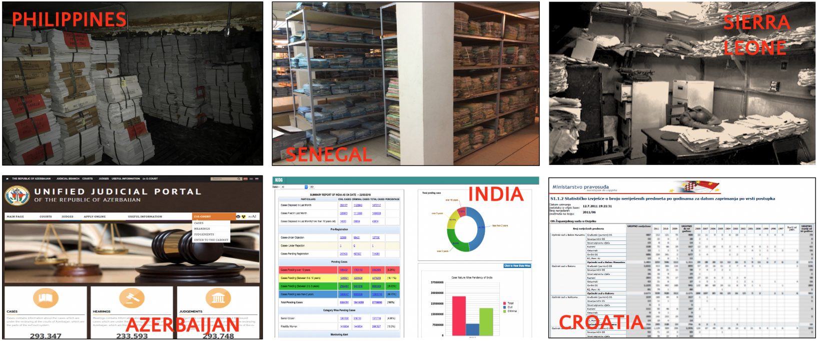
Would informing people about impacts of decisions improve intrinsic motivation?

AND IMPROVE POLICY

Data Revolution in Justice

- **AI + Economics => Law**
  - causal impacts of decisions
  - causal impacts of texts
  - **causal impacts of institutions**

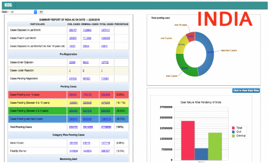
Data Revolution in Justice

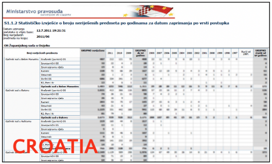


- Recent innovations have opened up new opportunities for delivery of justice
  - Increasingly digitized large-scale datasets
  - ML applications to produce interpretable data from unstructured text
  - Predictive models of decision-making to better understand biases and address them with digital interfaces

Ramos-Maqueda and Chen, World Development 2024

Data Revolution in Justice

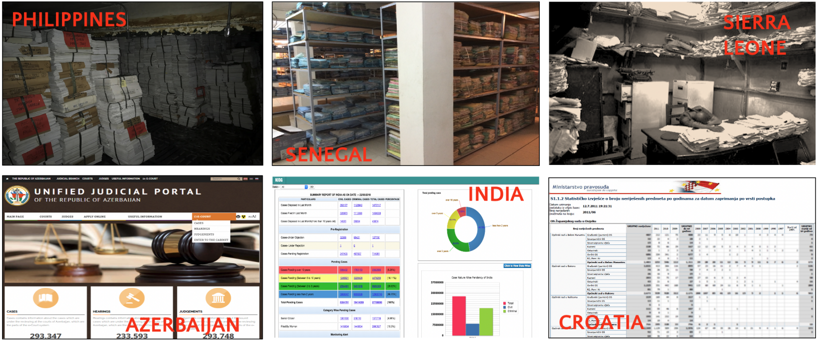




- Recent innovations have opened up new opportunities for delivery of justice
  - Increasingly digitized large-scale datasets
  - ML applications to produce interpretable data from unstructured text
  - Predictive models of decision-making to better understand biases and address them with digital interfaces

Ramos-Maqueda and Chen, World Development 2024

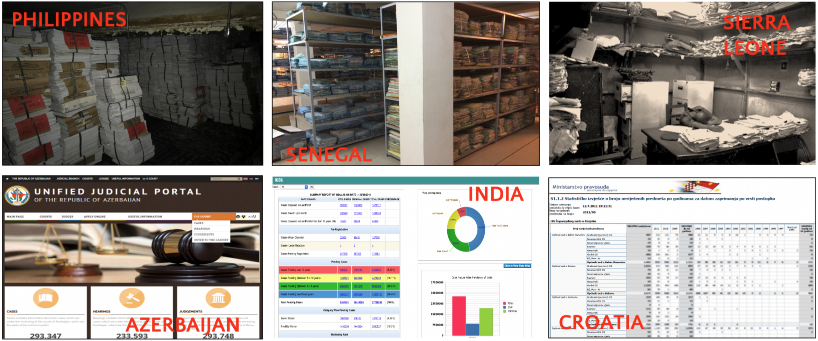
Data Revolution in Justice



- Recent innovations have opened up new opportunities for delivery of justice
  - Increasingly digitized large-scale datasets
  - ML applications to produce interpretable data from unstructured text
  - Predictive models of decision-making to better understand biases and address them with digital interfaces

Ramos-Maqueda and Chen, World Development 2024

Data Revolution in Justice



- Recent innovations have opened up new opportunities for delivery of justice
  - Increasingly digitized large-scale datasets
  - ML applications to produce interpretable data from unstructured text
  - Predictive models of decision-making to better understand biases and address them with digital interfaces

Ramos-Maqueda and Chen, World Development 2024

Recommending Actions to Each Other

Conciliator App | About | National Analysis | Conciliators | Legal Aid Centers | Login

All Centers
Geospatial Analysis
Temporal Analysis
Type of Case

Download CSV

Download PDF

Year of
2014 2020
2014 2019 2020

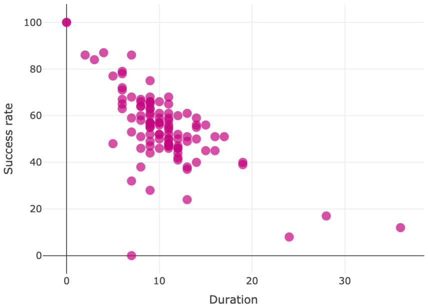
X: Duration

Y: Success rate

Show 25 entries

Search:

	Legal Aid Center	Number of cases	Success rate (%)	Duration of cases (days)	Avg. monthly caseload
1	CERCADO DE LIMA	2950	39	19	36
2	CHINCHA ALTA	972	47	12	19
3	MUNICIPALIDAD PROVINCIAL DE SATIPO	951	64	8	19
4	ICA	3648	60	9	36
5	AREQUIPA	1110	67	8	21
6	AMBO	85	58	11	8
7	HUANCAYO	1232	71	6	23



oTreeJustice

OR, AS MANAGEMENT TOOL, OBSERVING REGRESSIONS THAT THEY RUN

Recommending Actions to Each Other

Conciliator App | About | National Analysis ▾ | Conciliators ▾ | Legal Aid Centers ▾ | Login

National Analysis: All Centers · Geospatial Analysis · Temporal Analysis · Type of Case

Download CSV

Download PDF

Year of [...] 2014 – 2020

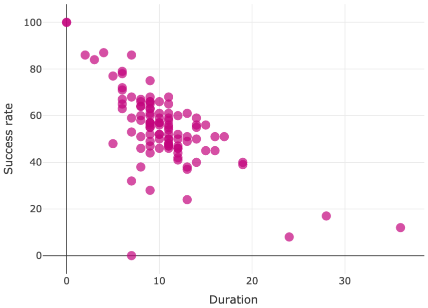
X: Duration

Y: Success rate

Show 25 entries

Search:

	Legal Aid Center	Number of cases	Success rate (%)	Duration of cases (days)	Avg. monthly caseload
1	CERCADO DE LIMA	2950	39	19	36
2	CHINCHA ALTA	972	47	12	19
3	MUNICIPALIDAD PROVINCIAL DE SATIPO	951	64	8	19
4	ICA	3648	60	9	36
5	AREQUIPA	1110	67	8	21
6	AMBO	85	58	11	8
7	HUANCAYO	1232	71	6	23

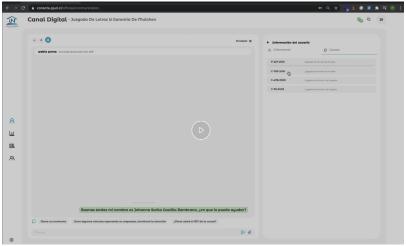


oTreeJustice

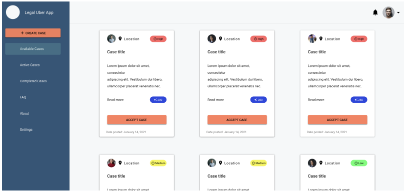
OR, AS MANAGEMENT TOOL, OBSERVING REGRESSIONS THAT THEY RUN

E-Justice Innovations

WhatsApp access to virtual courts

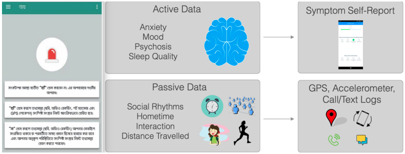


Uber-ization of case backlog



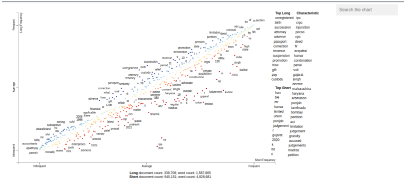
Chen, He, Mbau, J or Mechanism and Institution Design

Apps for missing cases



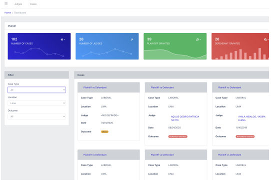
Ahmed, Ani, Alvi, Chen, Mbau, Saheel, Wahhaj

Open access legal search engines

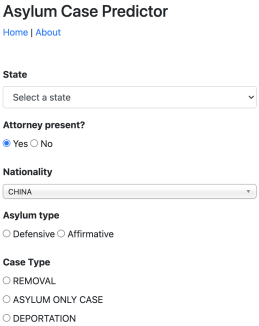


Human-Centric

Personalized case-based teaching



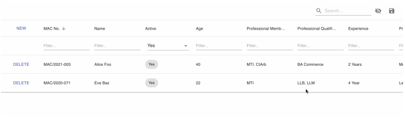
Predicted self



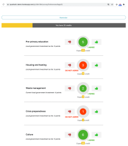
Building Capacity

Open source no-code tools for

Data entry and decision-support

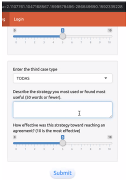


Understanding justice needs



Bassetti, Chen, Das, Dias, Mortoni, AI Magazine, NeurIPS 2021 (AI for Credible Elections)

Learning best practices



Increasing recognition-respect

Chen, Schonger, and Wickens
J Behavioral & Experimental Finance 2016

Interacting with AI at Work: Perceptions from the UK Judiciary, Solovey, Flanagan, Chen, CHIWORK '25

The Role of Justice in Development

A 20% decrease in case duration is associated with a 10% increase in GDP per capita
(Penn World tables)

Ramos-Maqueda and Chen, J of Economic Surveys R&R

The Role of Justice in Development

A 20% decrease in case duration is associated with a 10% increase in GDP per capita
(Penn World tables)

Ramos-Maqueda and Chen, J of Economic Surveys R&R

Data Science for Justice

Evidence from a Nationwide Randomized Experiment in Kenya

- Nation-wide experiment using the first digitized daily court records
  - Developed an algorithm to identify the greatest source of court delays
  - T1: provide actionable information (interpretable ML)
  - T2: + accountability (citizen groups)
  - Control: status quo (no information), 4-yr RCT across all 124 court stations

Data Science for Justice

Evidence from a Nationwide Randomized Experiment in Kenya

- Nation-wide experiment using the first digitized daily court records
  - Developed an algorithm to identify the greatest source of court delays
  - T1: provide actionable information (interpretable ML)
  - T2: + accountability (citizen groups)
  - Control: status quo (no information), 4-yr RCT across all 124 court stations

Data Science for Justice

Evidence from a Nationwide Randomized Experiment in Kenya

- Nation-wide experiment using the first digitized daily court records
  - Developed an algorithm to identify the greatest source of court delays
  - T1: provide actionable information (interpretable ML)
  - T2: + accountability (citizen groups)
  - Control: status quo (no information), 4-yr RCT across all 124 court stations

Data Science for Justice

Evidence from a Nationwide Randomized Experiment in Kenya

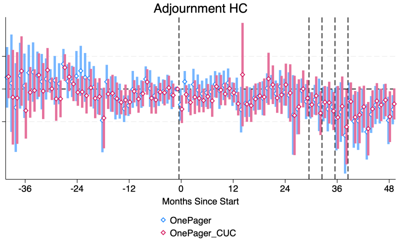
- Nation-wide experiment using the first digitized daily court records
  - Developed an algorithm to identify the greatest source of court delays
  - T1: provide actionable information (interpretable ML)
  - T2: + accountability (citizen groups)
  - Control: status quo (no information), 4-yr RCT across all 124 court stations

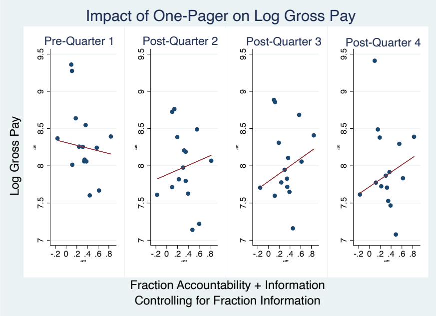
Data Science for Justice

Evidence from a Nationwide Randomized Experiment in Kenya

- Nation-wide experiment using the first digitized daily court records
  - Developed an algorithm to identify the greatest source of court delays
  - T1: provide actionable information (interpretable ML)
  - T2: + accountability (citizen groups)
  - Control: status quo (no information), 4-yr RCT across all 124 court stations

Accountability was key





20% decrease in case duration

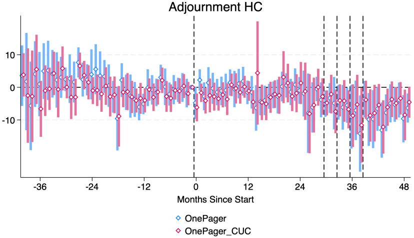
associated with 8% increased wages per capita

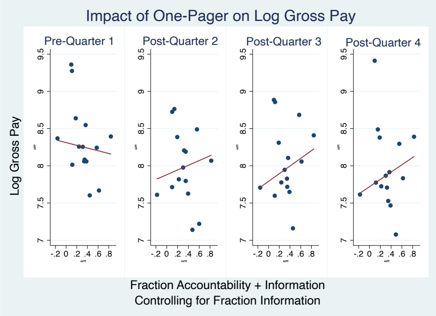
Chemin, Chen, Di Maro, Kimalu, Ramos-Maqueda

20% decrease in case duration after a court reform increased patenting by 6%

The Innovation Consequences of Judicial Efficiency, Kim, Shii, Verdi

Accountability was key





20% decrease in case duration

associated with 8% increased wages per capita

Chemin, Chen, Di Maro, Kimalu, Ramos-Maqueda

20% decrease in case duration after a court reform increased patenting by 6%

The Innovation Consequences of Judicial Efficiency, Kim, Shii, Verdi

Can digital platforms offering free legal information improve justice systems?

"bring knowledge of the law to the common people"

Keyword searches for automatic determination of most relevant clauses and judgments

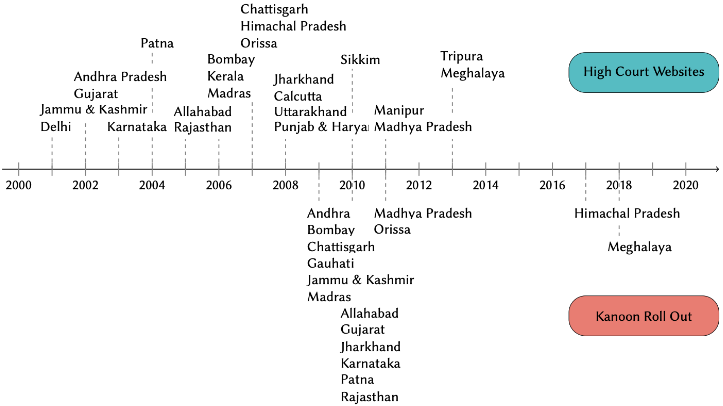


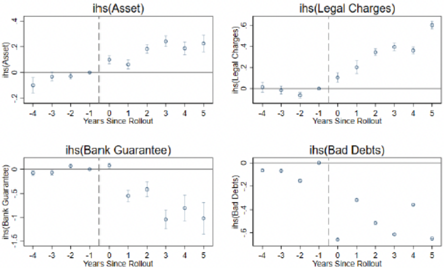
Figure: Roll Out Years for High Court Websites (top) and Kanoon (bottom)

Today, it is a "first-stop" for lawyers, 6 min per page, 2.9 M search queries and 1.5 M sessions per month

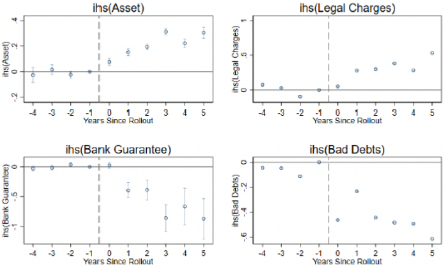
Event study analysis of firm financials

$$Y_{cfst} = \alpha + \sum_{j=2}^{4} \beta_j (lag_j)_{cfst} + \sum_{k=1}^{4} \gamma_k (lead_k)_{cfst} + \mu_s + \delta_f + \lambda_t + \epsilon_{cfst}$$

Firms with at least one case



General equilibrium



Sizeable impacts on assets and reduction of bad debt reinforce the findings of a 12% increase in employment in an RCT of free legal information to South African firms.

Bertrand and Crepon, Amer Econ J Applied 2021

Increased quantity & quality of entrepreneurship in terms of finances and innovation

Evidence from Mass Publicity of Chinese Court Decisions, Liu, Tian, Zhu

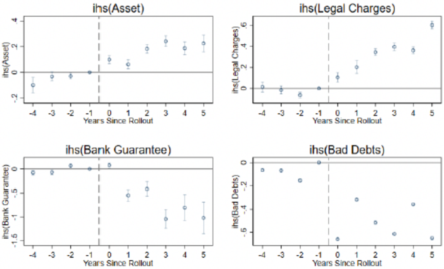
Highlights open source (common knowledge) for development, competition, & innovation

Bhupatiraju, Chen, Joshi, Neis, Journal of European Econ Assn, R&R

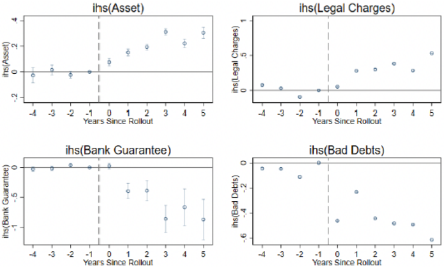
Event study analysis of firm financials

$$Y_{cfst} = \alpha + \sum_{j=2}^{4} \beta_j (lag_j)_{cfst} + \sum_{k=1}^{4} \gamma_k (lead_k)_{cfst} + \mu_s + \delta_f + \lambda_t + \epsilon_{cfst}$$

Firms with at least one case



General equilibrium



Sizeable impacts on assets and reduction of bad debt reinforce the findings of a 12% increase in employment in an RCT of free legal information to South African firms.

Bertrand and Crepon, Amer Econ J Applied 2021

Increased quantity & quality of entrepreneurship in terms of finances and innovation

Evidence from Mass Publicity of Chinese Court Decisions, Liu, Tian, Zhu

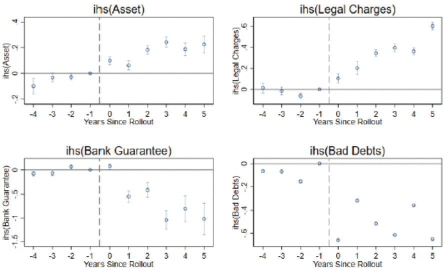
Highlights open source (common knowledge) for development, competition, & innovation

Bhupatiraju, Chen, Joshi, Neis, Journal of European Econ Assn, R&R

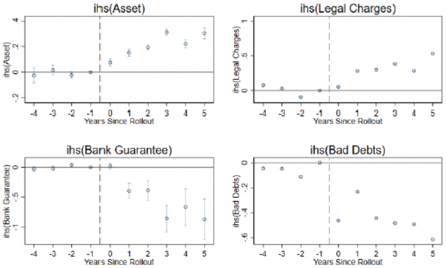
Event study analysis of firm financials

$$Y_{cfst} = \alpha + \sum_{j=2}^{4} \beta_j (lag_j)_{cfst} + \sum_{k=1}^{4} \gamma_k (lead_k)_{cfst} + \mu_s + \delta_f + \lambda_t + \epsilon_{cfst}$$

Firms with at least one case



General equilibrium



Sizeable impacts on assets and reduction of bad debt reinforce the findings of a 12% increase in employment in an RCT of free legal information to South African firms.

Bertrand and Crepon, Amer Econ J Applied 2021

Increased quantity & quality of entrepreneurship in terms of finances and innovation

Evidence from Mass Publicity of Chinese Court Decisions, Liu, Tian, Zhu

Highlights open source (common knowledge) for development, competition, & innovation

Bhupatiraju, Chen, Joshi, Neis, Journal of European Econ Assn, R&R

AI as Motor of Innovation

- Economics + Law => AI
- AI + Economics => Law
- AI + Law => Economics

AI as Motor of Innovation

- Economics + Law => AI
- AI + Economics => Law
- **AI + Law => Economics**

AI as Motor of Innovation

Future Directions:

- Fine-tuning with game behavior from human subjects for SelfGPT
- Is it predictive of real-world behavior? (current college choices, e.g. education/labor outcomes)
- Is it predictive of how they answer surveys today? (political attitudes, more games)
- At what age is SelfGPT predictive of what?
- What happens if we show the SelfGPT to the individual?
- Is SelfGPT or family/friends better at answering questions for the individual?

AI as Motor of Innovation

Future Directions:

- Fine-tuning with game behavior from human subjects for SelfGPT
- Is it predictive of real-world behavior? (current college choices, e.g. education/labor outcomes)
- Is it predictive of how they answer surveys today? (political attitudes, more games)
- At what age is SelfGPT predictive of what?
- What happens if we show the SelfGPT to the individual?
- Is SelfGPT or family/friends better at answering questions for the individual?

AI as Motor of Innovation

Future Directions:

- Fine-tuning with game behavior from human subjects for SelfGPT
- Is it predictive of real-world behavior? (current college choices, e.g. education/labor outcomes)
- Is it predictive of how they answer surveys today? (political attitudes, more games)
- At what age is SelfGPT predictive of what?
- What happens if we show the SelfGPT to the individual?
- Is SelfGPT or family/friends better at answering questions for the individual?

AI as Motor of Innovation

Future Directions:

- Fine-tuning with game behavior from human subjects for SelfGPT
- Is it predictive of real-world behavior? (current college choices, e.g. education/labor outcomes)
- Is it predictive of how they answer surveys today? (political attitudes, more games)
- At what age is SelfGPT predictive of what?
- What happens if we show the SelfGPT to the individual?
- Is SelfGPT or family/friends better at answering questions for the individual?

AI as Motor of Innovation

Future Directions:

- Fine-tuning with game behavior from human subjects for SelfGPT
- Is it predictive of real-world behavior? (current college choices, e.g. education/labor outcomes)
- Is it predictive of how they answer surveys today? (political attitudes, more games)
- At what age is SelfGPT predictive of what?
- What happens if we show the SelfGPT to the individual?
- Is SelfGPT or family/friends better at answering questions for the individual?

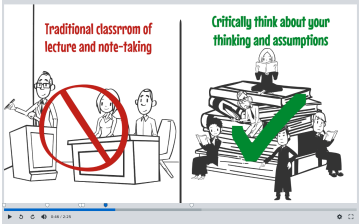
AI as Motor of Innovation

Future Directions:

- Fine-tuning with game behavior from human subjects for SelfGPT
- Is it predictive of real-world behavior? (current college choices, e.g. education/labor outcomes)
- Is it predictive of how they answer surveys today? (political attitudes, more games)
- At what age is SelfGPT predictive of what?
- What happens if we show the SelfGPT to the individual?
- Is SelfGPT or family/friends better at answering questions for the individual?

Field Experiments in Self-Reflection

Socratic training of judges

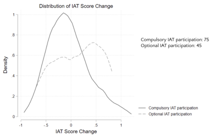


increased curiosity by 6.5 percentage points

	ITT			ToT		
VARIABLES	(1) VDO	(2) SBU	(3) Curiosity	(4) VDO	(5) SBU	(6) Curiosity
Assigned to Socratic	-0.016	0.028	0.065**			
	(0.042)	(0.047)	(0.027)			
	[0.904]	[0.896]	[0.030]			
Saw Socratic				-0.038	0.066	0.122**
				(0.100)	(0.109)	(0.053)
				[0.910]	[0.896]	[0.020]
Constant	0.980***	0.980***	0.874***	0.980***	0.980***	0.874***
	(0.030)	(0.029)	(0.023)	(0.030)	(0.029)	(0.023)
Observations	498	498	300	498	498	300
R-squared	0.000	0.001	0.013		0.004	
Individuals	498	498	300	498	498	300

Batistoni, Chen, and Silveira

option to self-reflect reduced bias



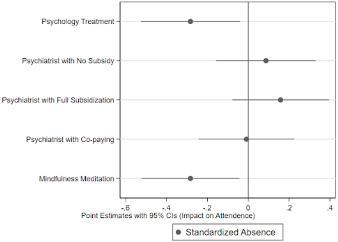
- [0, 0.15]: Low or none bias
-]0.15, 0.35]: Slight bias
-]0.35, 0.65]: Moderate bias
-]0.65, ...]: Strong bias

- Values **greater** than 0: Association between feminine and career

- Values **lower** than 0: Association between feminine and family

Batistoni, Chen, and Silveira

mindfulness improved civil servant performance



Mehmood, Ali, Chen, and Reinhart

AI and Mindfulness || Experimental History

- Predicted Self; Predicted Other for mindfulness and self-reflection

Community of Practice: Experimental Evidence from Peru's Judiciary, Chen, Ramos-Maqueda, and Silveira, J Public Economics R&R

- 1) train chatGPT on your own text to see if it predicts your own survey / experimental game responses
- 2) Different Predicted Selves

Reward-on-the-Line: Offline Reinforcement Learning Method for Conversational Agents, Chen, Lin, Wang, Yang, ACM AI, Ethics & Society 2025

- 3) elderly on ICU, living will, original intent of donors / legislators
- 4) do the same for the text of historical people, and study the behavioral economics/psychology of them
- 5) AI/ML to diagnose inattention

Judicial Inattention: Machine Prediction of Appeal Success, Chen and Zhang

- 6) Attention, Beliefs, and Expectations Shape Well-Being

Aungle, Langer, Chen, Matta

AI and Mindfulness || Experimental History

- Predicted Self; Predicted Other for mindfulness and self-reflection

Community of Practice: Experimental Evidence from Peru's Judiciary, Chen, Ramos-Maqueda, and Silveira, J Public Economics R&R

- 1) train chatGPT on your own text to see if it predicts your own survey / experimental game responses
- 2) Different Predicted Selves

Reward-on-the-Line: Offline Reinforcement Learning Method for Conversational Agents, Chen, Lin, Wang, Yang, ACM AI, Ethics & Society 2025

- 3) elderly on ICU, living will, original intent of donors / legislators
- 4) do the same for the text of historical people, and study the behavioral economics/psychology of them
- 5) AI/ML to diagnose inattention

Judicial Inattention: Machine Prediction of Appeal Success, Chen and Zhang

- 6) Attention, Beliefs, and Expectations Shape Well-Being

Aungle, Langer, Chen, Matta

AI and Mindfulness || Experimental History

- Predicted Self; Predicted Other for mindfulness and self-reflection

Community of Practice: Experimental Evidence from Peru's Judiciary, Chen, Ramos-Maqueda, and Silveira, J Public Economics R&R

- 1) train chatGPT on your own text to see if it predicts your own survey / experimental game responses
- 2) Different Predicted Selves

Reward-on-the-Line: Offline Reinforcement Learning Method for Conversational Agents, Chen, Lin, Wang, Yang, ACM AI, Ethics & Society 2025

- 3) elderly on ICU, living will, original intent of donors / legislators
- 4) do the same for the text of historical people, and study the behavioral economics/psychology of them
- 5) AI/ML to diagnose inattention

Judicial Inattention: Machine Prediction of Appeal Success, Chen and Zhang

- 6) Attention, Beliefs, and Expectations Shape Well-Being

Aungle, Langer, Chen, Matta

AI and Mindfulness || Experimental History

- Predicted Self; Predicted Other for mindfulness and self-reflection

Community of Practice: Experimental Evidence from Peru's Judiciary, Chen, Ramos-Maqueda, and Silveira, J Public Economics R&R

- 1) train chatGPT on your own text to see if it predicts your own survey / experimental game responses
- 2) Different Predicted Selves

Reward-on-the-Line: Offline Reinforcement Learning Method for Conversational Agents, Chen, Lin, Wang, Yang, ACM AI, Ethics & Society 2025

- 3) elderly on ICU, living will, original intent of donors / legislators
- 4) do the same for the text of historical people, and study the behavioral economics/psychology of them
- 5) AI/ML to diagnose inattention

Judicial Inattention: Machine Prediction of Appeal Success, Chen and Zhang

- 6) Attention, Beliefs, and Expectations Shape Well-Being

Aungle, Langer, Chen, Matta

AI and Mindfulness || Experimental History

- Predicted Self; Predicted Other for mindfulness and self-reflection

Community of Practice: Experimental Evidence from Peru's Judiciary, Chen, Ramos-Maqueda, and Silveira, J Public Economics R&R

- 1) train chatGPT on your own text to see if it predicts your own survey / experimental game responses
- 2) Different Predicted Selves

Reward-on-the-Line: Offline Reinforcement Learning Method for Conversational Agents, Chen, Lin, Wang, Yang, ACM AI, Ethics & Society 2025

- 3) elderly on ICU, living will, original intent of donors / legislators
- 4) do the same for the text of historical people, and study the behavioral economics/psychology of them
- 5) AI/ML to diagnose inattention

Judicial Inattention: Machine Prediction of Appeal Success, Chen and Zhang

- 6) Attention, Beliefs, and Expectations Shape Well-Being

Aungle, Langer, Chen, Matta

AI and Mindfulness || Experimental History

- Predicted Self; Predicted Other for mindfulness and self-reflection

Community of Practice: Experimental Evidence from Peru's Judiciary, Chen, Ramos-Maqueda, and Silveira, J Public Economics R&R

- 1) train chatGPT on your own text to see if it predicts your own survey / experimental game responses
- 2) Different Predicted Selves

Reward-on-the-Line: Offline Reinforcement Learning Method for Conversational Agents, Chen, Lin, Wang, Yang, ACM AI, Ethics & Society 2025

- 3) elderly on ICU, living will, original intent of donors / legislators
- 4) do the same for the text of historical people, and study the behavioral economics/psychology of them
- 5) AI/ML to diagnose inattention

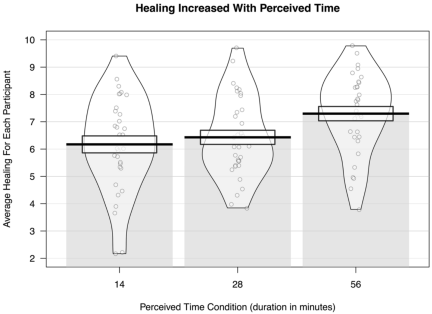
Judicial Inattention: Machine Prediction of Appeal Success, Chen and Zhang

- 6) Attention, Beliefs, and Expectations Shape Well-Being

Aungle, Langer, Chen, Matta

Subjective Perception of Time

- Wounds healed faster when participants believed more time had passed



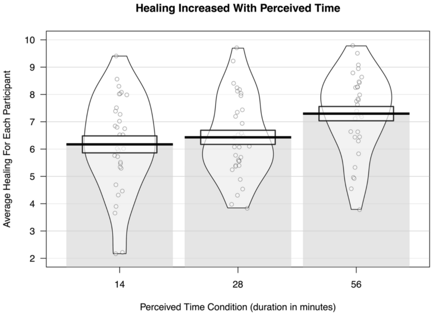
Physical healing as a function of perceived time, Aungle and Langer, Nature Sci Reports, 2023

- Neural correlates (EEG, fMRI) of fatigue match subjective perception of time

Mind Over Muscle, Matta, Glories, Alamia, Duclay, Psychophysiology, 2023

Subjective Perception of Time

- Wounds healed faster when participants believed more time had passed



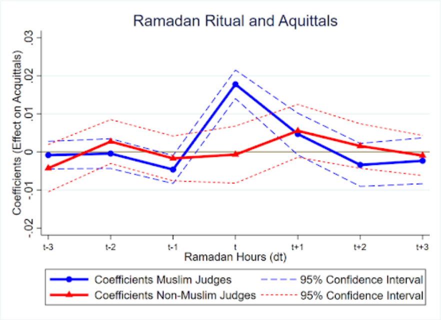
Physical healing as a function of perceived time, Aungle and Langer, Nature Sci Reports, 2023

- Neural correlates (EEG, fMRI) of fatigue match subjective perception of time

Mind Over Muscle, Matta, Glories, Alamia, Duclay, Psychophysiology, 2023

Formal Model of Interoception and Attention

Muslim judges are more lenient the longer is Ramadan



Pakistan and India

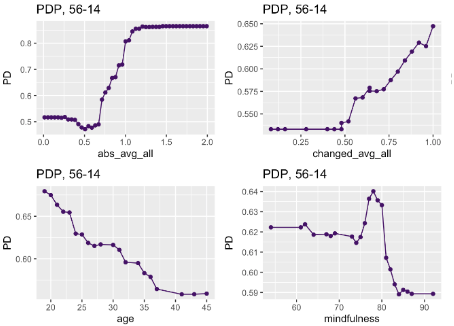
Mehmood, Seror, Chen, Nature Human Behavior 2023

formalizing "anthropological lit arguing that Ramadan fasting is associated with greater reflection and self-control"

connected to cognitive economics (attention & memory) / mindful econ *(J of Econ Persp 2016)*

Subjective Perception of Time

attention to healing variance drives healing effects



Mind-Body Economics: Formalizing Embodied Models of Health, Aungle, Chen, Loecher, and Matta

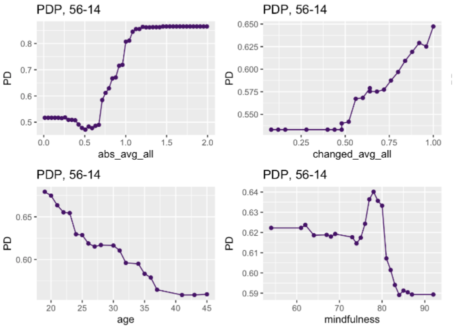
Manipulating Perceived Duration with Visual Augmentations, Harrison et al, CHI 2010

The Long and Short of It: Consumers Lose Track of Time, Amano, Nam, Shen, J of Marketing Research R/R

Would changing our sense of time reduce fatigue and error?

Subjective Perception of Time

attention to healing variance drives healing effects



Mind-Body Economics: Formalizing Embodied Models of Health, Aungle, Chen, Loecher, and Matta

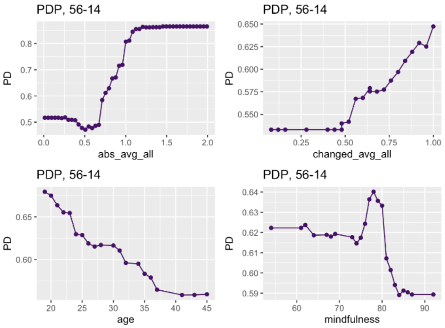
Manipulating Perceived Duration with Visual Augmentations, Harrison et al, CHI 2010

The Long and Short of It: Consumers Lose Track of Time, Amano, Nam, Shen, J of Marketing Research R/R

Would changing our sense of time reduce fatigue and error?

Subjective Perception of Time

attention to healing variance drives healing effects



Mind-Body Economics: Formalizing Embodied Models of Health, Aungle, Chen, Loecher, and Matta

Manipulating Perceived Duration with Visual Augmentations, Harrison et al, CHI 2010

The Long and Short of It: Consumers Lose Track of Time, Amano, Nam, Shen, J of Marketing Research R/R

Would changing our sense of time reduce fatigue and error?

Human-Centered AI

Can AI improve decision-making? *Babic, Chen, Evgeniou, and Fayard, Harvard Business Review 2020*

- Cognitive science and psychology suggests that humans have limited and imperfect reasoning capacities (Tversky and Kahneman 1986; Eyster 2019)
- Gambler's fallacy, mood, time of day, order, ...

  Chen, Moskowitz, and Shue, Quarterly J Econ 2016

  - highlight fragility of courts
    - "In a crowded immigration court, 7 minutes to decide a family's future" (Wash Post 2/2/14)
- Policy discussion tends to revolve around having AI replace humans or suggest the optimal decision
- Consider instead an incremental approach based on Enlightenment and Romantic ideals of the self: self-knowledge, self-expression

(Charles Taylor, Sources of the Self, 1989; The Ethics of Authenticity, 1992)

Human-Centered AI

Can AI improve decision-making? *Babic, Chen, Evgeniou, and Fayard, Harvard Business Review 2020*

- Cognitive science and psychology suggests that humans have limited and imperfect reasoning capacities (Tversky and Kahneman 1986; Eyster 2019)

- Gambler's fallacy, mood, time of day, order, ...

  Chen, Moskowitz, and Shue, Quarterly J Econ 2016

  - highlight fragility of courts
    - ★ "In a crowded immigration court, 7 minutes to decide a family's future" (Wash Post 2/2/14)

- Policy discussion tends to revolve around having AI replace humans or suggest the optimal decision

- Consider instead an incremental approach based on Enlightenment and Romantic ideals of the self: self-knowledge, self-expression

(Charles Taylor, Sources of the Self, 1989; The Ethics of Authenticity, 1992)

Asian J of Law and Economics 2023

Human-Centered AI

Can AI improve decision-making? *Babic, Chen, Evgeniou, and Fayard, Harvard Business Review 2020*

- Cognitive science and psychology suggests that humans have limited and imperfect reasoning capacities (Tversky and Kahneman 1986; Eyster 2019)
- Gambler's fallacy, mood, time of day, order, ...

  Chen, Moskowitz, and Shue, Quarterly J Econ 2016

  - highlight fragility of courts
    - "In a crowded immigration court, 7 minutes to decide a family's future" (Wash Post 2/2/14)
- Policy discussion tends to revolve around having AI replace humans or suggest the optimal decision
- Consider instead an incremental approach based on Enlightenment and Romantic ideals of the self: self-knowledge, self-expression

(Charles Taylor, Sources of the Self, 1989; The Ethics of Authenticity, 1992)

Asian J of Law and Economics 2023

Human-Centered AI

Can AI improve decision-making? *Babic, Chen, Evgeniou, and Fayard, Harvard Business Review 2020*

- Cognitive science and psychology suggests that humans have limited and imperfect reasoning capacities (Tversky and Kahneman 1986; Eyster 2019)

- Gambler's fallacy, mood, time of day, order, ...

  Chen, Moskowitz, and Shue, Quarterly J Econ 2016

  - highlight fragility of courts
    - "In a crowded immigration court, 7 minutes to decide a family's future" (Wash Post 2/2/14)

- Policy discussion tends to revolve around having AI replace humans or suggest the optimal decision

- Consider instead an incremental approach based on Enlightenment and Romantic ideals of the self: self-knowledge, self-expression

(Charles Taylor, Sources of the Self, 1989; The Ethics of Authenticity, 1992)

Asian J of Law and Economics 2023

Human-Centered AI

Can AI improve decision-making? *Babic, Chen, Evgeniou, and Fayard, Harvard Business Review 2020*

- Cognitive science and psychology suggests that humans have limited and imperfect reasoning capacities (Tversky and Kahneman 1986; Eyster 2019)
- Gambler's fallacy, mood, time of day, order, ...

  Chen, Moskowitz, and Shue, Quarterly J Econ 2016
  - highlight fragility of courts
    - "In a crowded immigration court, 7 minutes to decide a family's future" (Wash Post 2/2/14)
- Policy discussion tends to revolve around having AI replace humans or suggest the optimal decision
- Consider instead an incremental approach based on Enlightenment and Romantic ideals of the self: self-knowledge, self-expression

(Charles Taylor, Sources of the Self, 1989; The Ethics of Authenticity, 1992)

Asian J of Law and Economics 2023

Human-Centered AI and Self-Reflection

One of the most remarkable aspects of this self-evolution is the emergence of sophisticated behaviors as the test-time computation increases. Behaviors such as reflection—where the model revisits and reevaluates its previous steps—and the exploration of alternative approaches to problem-solving arise spontaneously. These behaviors are not explicitly programmed but instead emerge as a result of the model's interaction with the reinforcement learning environment. This spontaneous development significantly enhances DeepSeek-R1-Zero's reasoning capabilities, enabling it to tackle more challenging tasks with greater efficiency and accuracy.

DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning

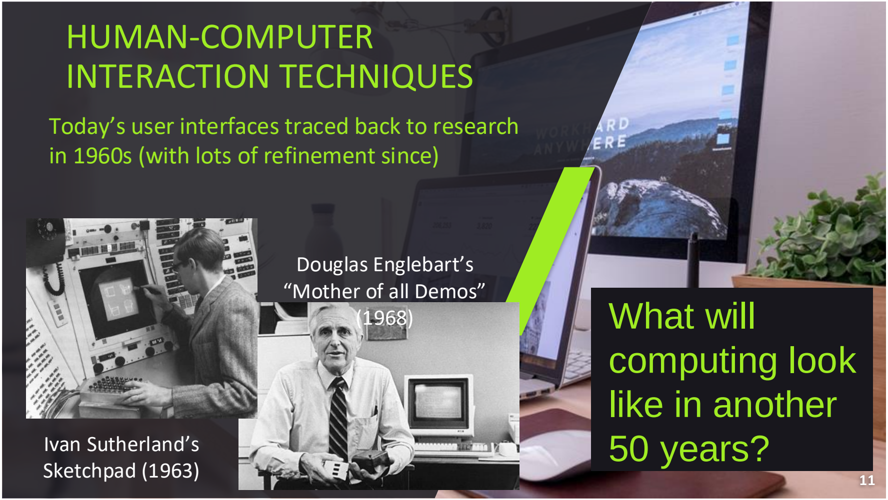
Human-Centered AI and Self-Reflection

One of the most remarkable aspects of this self-evolution is the emergence of sophisticated behaviors as the test-time computation increases. Behaviors such as reflection—where the model revisits and reevaluates its previous steps—and the exploration of alternative approaches to problem-solving arise spontaneously. These behaviors are not explicitly programmed but instead emerge as a result of the model's interaction with the reinforcement learning environment. This spontaneous development significantly enhances DeepSeek-R1-Zero's reasoning capabilities, enabling it to tackle more challenging tasks with greater efficiency and accuracy.

DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning

Human-Centered AI Wearables

(slides credit Radcliffe: Erin Solovey)



fNIRS for Measuring Attention

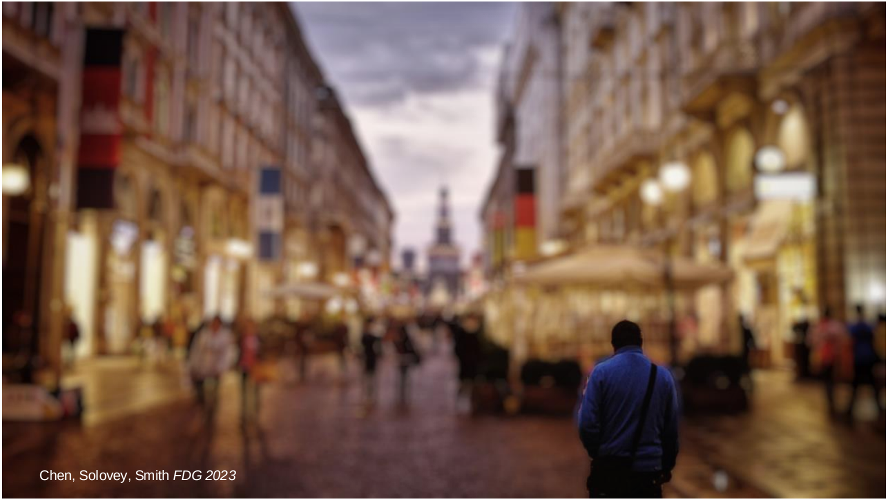
Functional near-infrared spectroscopy

Brain as **implicit, supplementary** input

- Augment traditional input devices
- Wider group of users, beyond disabled
- Passive, implicit input channel
- Capture subtle cognitive state changes
- Input to adaptive interactive system
- Real-time, continuous data

digital yarmulke and headbands

fNIRS for Measuring Attention



Chen, Solovey, Smith *FDG 2023*

fNIRS for Measuring Attention

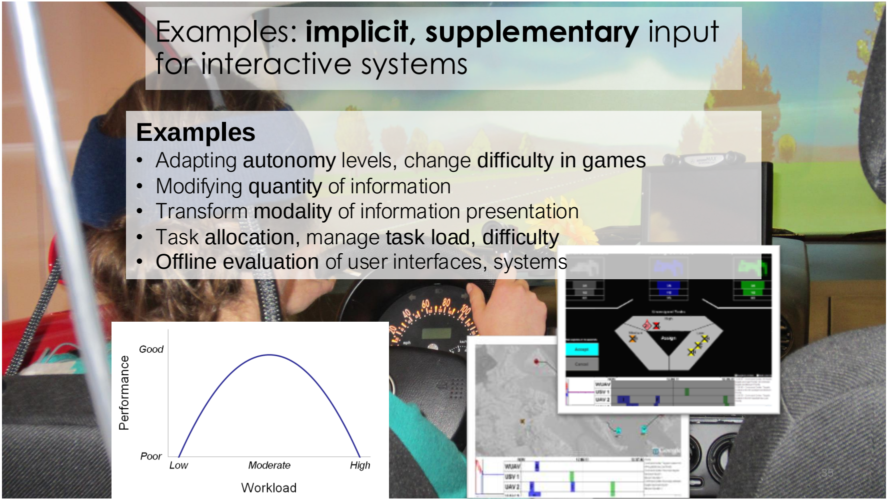
Chen, Solovey, Smith *FDG 2023*

fNIRS for Measuring Attention

Examples: **implicit, supplementary** input for interactive systems

Examples

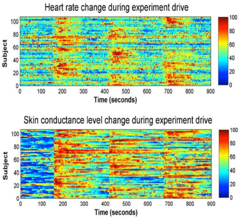
- Adapting **autonomy** levels, change **difficulty in games**
- Modifying **quantity** of information
- Transform **modality** of information presentation
- Task **allocation**, manage **task load, difficulty**
- **Offline evaluation** of user interfaces, systems



fNIRS for Measuring Attention

Judge thinking caps

Example: Classifying Driver Workload Using Physiological & Driving Performance



- Classify cognitive workload level
- Body sensor & task data as input
- Real-world task, large field studies

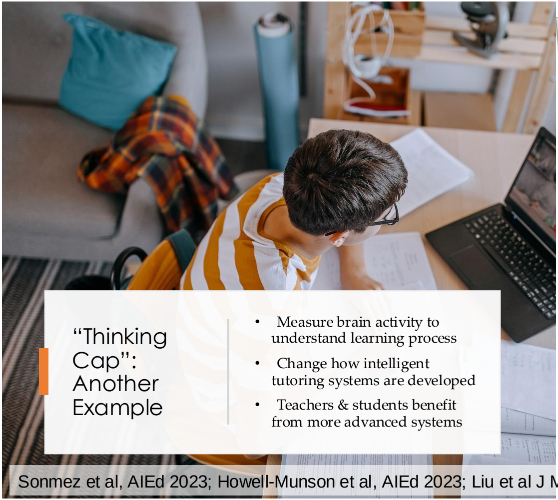
Solovey, et al., (2014) Classifying driver workload using physiological and driving performance data: two field studies. In Proc. *CHI '14.*, 4057–4066.

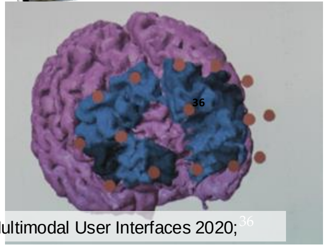
reduce error / enhance recognition-respect

fNIRS for Measuring Learning

Judge learning caps

thinking hard vs. mind-wandering





Sonmez et al, AIEd 2023; Howell-Munson et al, AIEd 2023; Liu et al J Multimodal User Interfaces 2020; 36

actively encoding or accessing long-term memory

fNIRS for Measuring Coordination

Judicial panels

Distributed teamwork with AI in critical settings

Can brain and physiological data provide insights that could improve team performance?

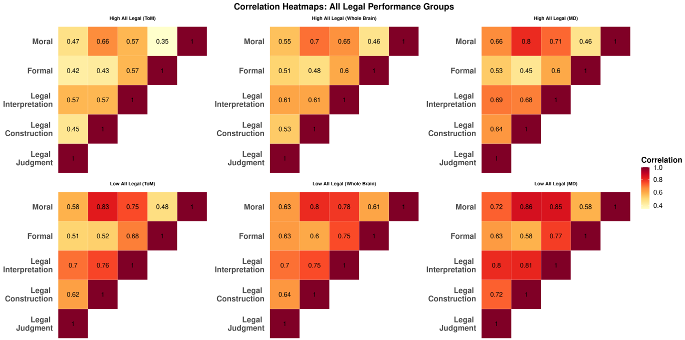
Could we use these signals to detect when someone is experiencing a critical state (e.g. excessive workload, distraction, focus) and effectively improve team performance?

How do team members perceive and adjust to awareness of cognitive states across teams?

What does it mean when we begin to sense aspects of our teammates' cognitive states that had not been previously accessible to us?

Neural Correlates of Legal Reasoning Performance

An fMRI Analysis



High scorers separate moral judgment, formal logic, and contract interpretation into distinct processing streams
in domain-general control regions—fronto-parietal MD—and in social-cognition regions—TPJ and medial PFC

Effective legal thinkers compartmentalize these cognitive operations neurally

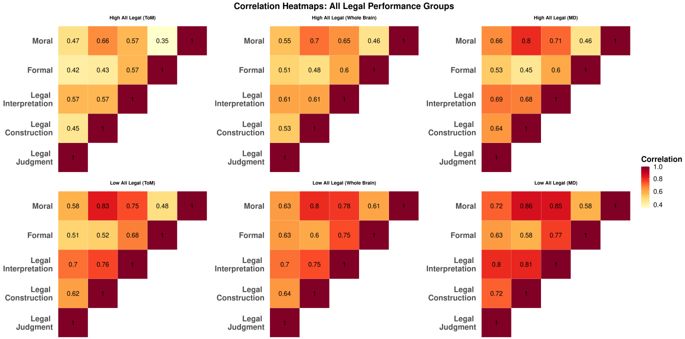
Martinez, Chen, Solovey, Ivanova

MIGHT FOUNDATION MODELS FOR fNIRS DATA IN (MORAL) DECISION-MAKING HAVE THE SAME ROLE AS LLMs?

HOW SHOULD WE FOSTER AI INNOVATION?

Neural Correlates of Legal Reasoning Performance

An fMRI Analysis



High scorers separate moral judgment, formal logic, and contract interpretation into distinct processing streams
in domain-general control regions—fronto-parietal MD—and in social-cognition regions—TPJ and medial PFC

Effective legal thinkers compartmentalize these cognitive operations neurally

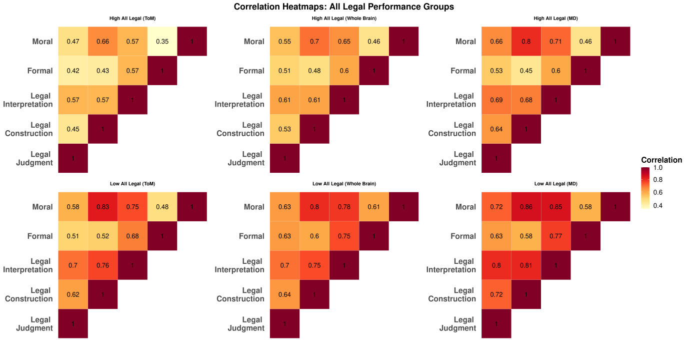
Martinez, Chen, Solovey, Ivanova

might foundation models for fNIRS data in (moral) decision-making have the same role as LLMs?

how should we foster AI innovation?

Neural Correlates of Legal Reasoning Performance

An fMRI Analysis



High scorers separate moral judgment, formal logic, and contract interpretation into distinct processing streams
in domain-general control regions—fronto-parietal MD—and in social-cognition regions—TPJ and medial PFC

Effective legal thinkers compartmentalize these cognitive operations neurally

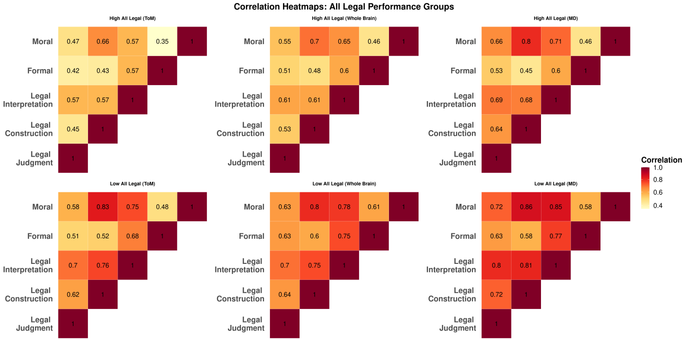
Martinez, Chen, Solovey, Ivanova

MIGHT FOUNDATION MODELS FOR fNIRS DATA IN (MORAL) DECISION-MAKING HAVE THE SAME ROLE AS LLMs?

HOW SHOULD WE FOSTER AI INNOVATION?

Neural Correlates of Legal Reasoning Performance

An fMRI Analysis



High scorers separate moral judgment, formal logic, and contract interpretation into distinct processing streams

in domain-general control regions—fronto-parietal MD—and in social-cognition regions—TPJ and medial PFC

Effective legal thinkers compartmentalize these cognitive operations neurally

Martinez, Chen, Solovey, Ivanova

MIGHT FOUNDATION MODELS FOR fNIRS DATA IN (MORAL) DECISION-MAKING HAVE THE SAME ROLE AS LLMs?

HOW SHOULD WE FOSTER AI INNOVATION?

Precision Medicine

How to Regulate and Finance N-of-1 Precision Medicine

- Randomized trials impossible with a sample size of one.
- Medical-malpractice and products-liability suits collapse on causation and "customary care"
- Ultra-small markets erase the profit motive
- Every bespoke therapy generates data that reduces the cost of the next one—a positive externality standard IP law can't capture.
- Multi-sided platform: labs, patients/payers, regulators *(Rochet and Tirole, Rand 2006)*
  - **Subsidies**: Payers fund per-case "data bounties"; labs receive negative fees to publish manufacturing & outcome data.
  - **Disclosure**: 30-day raw-read uploads + pre-registered analytics = real-time learning across centers.
  - **Safe-Harbor Liability**: Compliance routes injuries through a no-fault fund, preserving deterrence without bankrupting tiny labs.
- Policy Payoff
  - Turns one-off treatments into a learning network—scaling precision medicine.

Precision Medicine

How to Regulate and Finance N-of-1 Precision Medicine

- Randomized trials impossible with a sample size of one.
- Medical-malpractice and products-liability suits collapse on causation and "customary care"
- Ultra-small markets erase the profit motive
- Every bespoke therapy generates data that reduces the cost of the next one—a positive externality standard IP law can't capture.
- Multi-sided platform: labs, patients/payers, regulators *(Rochet and Tirole, Rand 2006)*
  - **Subsidies**: Payers fund per-case "data bounties"; labs receive negative fees to publish manufacturing & outcome data.
  - **Disclosure**: 30-day raw-read uploads + pre-registered analytics = real-time learning across centers.
  - **Safe-Harbor Liability**: Compliance routes injuries through a no-fault fund, preserving deterrence without bankrupting tiny labs.
- Policy Payoff
  - Turns one-off treatments into a learning network—scaling precision medicine.

Precision Medicine

How to Regulate and Finance N-of-1 Precision Medicine

- Randomized trials impossible with a sample size of one.
- Medical-malpractice and products-liability suits collapse on causation and "customary care"
- Ultra-small markets erase the profit motive
- Every bespoke therapy generates data that reduces the cost of the next one—a positive externality standard IP law can't capture.
- Multi-sided platform: labs, patients/payers, regulators *(Rochet and Tirole, Rand 2006)*
  - **Subsidies**: Payers fund per-case "data bounties"; labs receive negative fees to publish manufacturing & outcome data.
  - **Disclosure**: 30-day raw-read uploads + pre-registered analytics = real-time learning across centers.
  - **Safe-Harbor Liability**: Compliance routes injuries through a no-fault fund, preserving deterrence without bankrupting tiny labs.
- Policy Payoff
  - Turns one-off treatments into a learning network—scaling precision medicine.

Precision Medicine

How to Regulate and Finance N-of-1 Precision Medicine

- Randomized trials impossible with a sample size of one.
- Medical-malpractice and products-liability suits collapse on causation and "customary care"
- Ultra-small markets erase the profit motive
- Every bespoke therapy generates data that reduces the cost of the next one—a positive externality standard IP law can't capture.
- Multi-sided platform: labs, patients/payers, regulators *(Rochet and Tirole, Rand 2006)*
  - **Subsidies**: Payers fund per-case "data bounties"; labs receive negative fees to publish manufacturing & outcome data.
  - **Disclosure**: 30-day raw-read uploads + pre-registered analytics = real-time learning across centers.
  - **Safe-Harbor Liability**: Compliance routes injuries through a no-fault fund, preserving deterrence without bankrupting tiny labs.
- Policy Payoff
  - Turns one-off treatments into a learning network—scaling precision medicine.

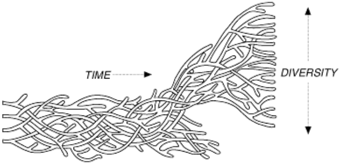
Precision Medicine

How to Regulate and Finance N-of-1 Precision Medicine

- Randomized trials impossible with a sample size of one.
- Medical-malpractice and products-liability suits collapse on causation and "customary care"
- Ultra-small markets erase the profit motive
- Every bespoke therapy generates data that reduces the cost of the next one—a positive externality standard IP law can't capture.
- Multi-sided platform: labs, patients/payers, regulators *(Rochet and Tirole, Rand 2006)*
  - **Subsidies**: Payers fund per-case "data bounties"; labs receive negative fees to publish manufacturing & outcome data.
  - **Disclosure**: 30-day raw-read uploads + pre-registered analytics = real-time learning across centers.
  - **Safe-Harbor Liability**: Compliance routes injuries through a no-fault fund, preserving deterrence without bankrupting tiny labs.
- Policy Payoff
  - Turns one-off treatments into a learning network—scaling precision medicine.

Measuring Innovation

- limitations of patent data
  - *"Inventions may be the wrong unit of measurement... and may be a misleading quantum"* [Griliches 1962]
  - Many inventions are not patented [Cohen et al. 2000, Levin et al. 1987]
  - Patents often filed strategically [Hall and Harhoff 2012, Harhoff 2016]
- Scientific publications



How Does Science Progress? Chen and Parsa

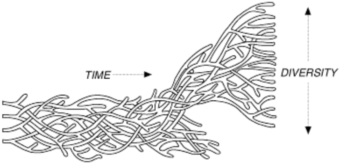
Recombinant Growth, Weitzman, Quarterly J Econ 1998

Knowledge Recombination, Gruber, Harhoff, Hoisl, Management Science 2013

- the measurement of ideas, the spread of ideas, paradigm shifts in the judicial space can be explored in other text-based spaces

Measuring Innovation

- limitations of patent data
  - *"Inventions may be the wrong unit of measurement... and may be a misleading quantum"* [Griliches 1962]
  - Many inventions are not patented [Cohen et al. 2000, Levin et al. 1987]
  - Patents often filed strategically [Hall and Harhoff 2012, Harhoff 2016]
- Scientific publications



How Does Science Progress? Chen and Parsa

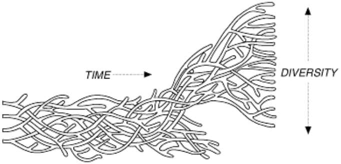
Recombinant Growth, Weitzman, Quarterly J Econ 1998

Knowledge Recombination, Gruber, Harhoff, Hoisl, Management Science 2013

- the measurement of ideas, the spread of ideas, paradigm shifts in the judicial space can be explored in other text-based spaces

Measuring Innovation

- limitations of patent data
  - *"Inventions may be the wrong unit of measurement... and may be a misleading quantum"* [Griliches 1962]
  - Many inventions are not patented [Cohen et al. 2000, Levin et al. 1987]
  - Patents often filed strategically [Hall and Harhoff 2012, Harhoff 2016]
- Scientific publications



How Does Science Progress? Chen and Parsa

Recombinant Growth, Weitzman, Quarterly J Econ 1998

Knowledge Recombination, Gruber, Harhoff, Hoisl, Management Science 2013

- the measurement of ideas, the spread of ideas, paradigm shifts in the judicial space can be explored in other text-based spaces

Fostering Innovation

- legal doctrine, security of property rights, and trust in the law
  - how these can stimulate investment, reliance, growth
  - these are standard ideas that also work in ideas creation
- Fiscal Policy:
  - summary of the literature finds that a 10 percent fall in the tax price of R&D generates a 10 percent increase in R&D in the long-run

    Bloom, Van Reenen, and Williams, Journal of Econ Perspectives 2019
  - alternatives, e.g., prizes, contests, patent buyouts, open access
- Social Policy: individual determinants of innovation

  Bell, Chetty, Jaravel, Petkova, Van Reenan, Quarterly J Econ 2019
  - intrinsic motivation (social emotional learning measurements)

    Terrier, Chen, and Sutter, PNAS 2021
  - ed tech records and administrative data

    Chen, Ertac, Evgeniou, Nadaf, Miao, Yilmax, Nature Education 2024

AI Education as State Capacity: Experimental Evidence from Pakistan, Mehmood, Naseer, and Chen, J Development Econ R&R

The Effects of Entrepreneurship Education, Von Graevenitz, Harhoff, Weber, J Econ Behavioral Org 2010

Fostering Innovation

- legal doctrine, security of property rights, and trust in the law
  - how these can stimulate investment, reliance, growth
  - these are standard ideas that also work in ideas creation
- **Fiscal Policy**:
  - summary of the literature finds that a 10 percent fall in the tax price of R&D generates a 10 percent increase in R&D in the long-run

    Bloom, Van Reenen, and Williams, Journal of Econ Perspectives 2019
  - alternatives, e.g., prizes, contests, patent buyouts, open access
- **Social Policy**: individual determinants of innovation

  Bell, Chetty, Jaravel, Petkova, Van Reenan, Quarterly J Econ 2019
  - intrinsic motivation (social emotional learning measurements)

    Terrier, Chen, and Sutter, PNAS 2021
  - ed tech records and administrative data

    Chen, Ertac, Evgeniou, Nadaf, Miao, Yilmax, Nature Education 2024

AI Education as State Capacity: Experimental Evidence from Pakistan, Mehmood, Naseer, and Chen, J Development Econ R&R

The Effects of Entrepreneurship Education, Von Graevenitz, Harhoff, Weber, J Econ Behavioral Org 2010

Fostering Innovation

- legal doctrine, security of property rights, and trust in the law
  - how these can stimulate investment, reliance, growth
  - these are standard ideas that also work in ideas creation
- **Fiscal Policy**:
  - summary of the literature finds that a 10 percent fall in the tax price of R&D generates a 10 percent increase in R&D in the long-run

    Bloom, Van Reenen, and Williams, Journal of Econ Perspectives 2019
  - alternatives, e.g., prizes, contests, patent buyouts, open access
- **Social Policy**: individual determinants of innovation

  Bell, Chetty, Jaravel, Petkova, Van Reenan, Quarterly J Econ 2019
  - intrinsic motivation (social emotional learning measurements)

    Terrier, Chen, and Sutter, PNAS 2021
  - ed tech records and administrative data

    Chen, Ertac, Evgeniou, Nadaf, Miao, Yilmax, Nature Education 2024

AI Education as State Capacity: Experimental Evidence from Pakistan, Mehmood, Naseer, and Chen, J Development Econ R&R

The Effects of Entrepreneurship Education, Von Graevenitz, Harhoff, Weber, J Econ Behavioral Org 2010

Fostering Innovation

- legal doctrine, security of property rights, and trust in the law
  - how these can stimulate investment, reliance, growth
  - these are standard ideas that also work in ideas creation
- **Fiscal Policy**:
  - summary of the literature finds that a 10 percent fall in the tax price of R&D generates a 10 percent increase in R&D in the long-run

    Bloom, Van Reenen, and Williams, Journal of Econ Perspectives 2019
  - alternatives, e.g., prizes, contests, patent buyouts, open access
- **Social Policy**: individual determinants of innovation

  Bell, Chetty, Jaravel, Petkova, Van Reenan, Quarterly J Econ 2019
  - intrinsic motivation (social emotional learning measurements)

    Terrier, Chen, and Sutter, PNAS 2021
  - ed tech records and administrative data

    Chen, Ertac, Evgeniou, Nadaf, Miao, Yilmax, Nature Education 2024

AI Education as State Capacity: Experimental Evidence from Pakistan, Mehmood, Naseer, and Chen, J Development Econ R&R

The Effects of Entrepreneurship Education, Von Graevenitz, Harhoff, Weber, J Econ Behavioral Org 2010

Fostering Innovation

- legal doctrine, security of property rights, and trust in the law
  - how these can stimulate investment, reliance, growth
  - these are standard ideas that also work in ideas creation
- **Fiscal Policy**:
  - summary of the literature finds that a 10 percent fall in the tax price of R&D generates a 10 percent increase in R&D in the long-run

    Bloom, Van Reenen, and Williams, Journal of Econ Perspectives 2019
  - alternatives, e.g., prizes, contests, patent buyouts, open access
- **Social Policy**: individual determinants of innovation

  Bell, Chetty, Jaravel, Petkova, Van Reenan, Quarterly J Econ 2019
  - intrinsic motivation (social emotional learning measurements)

    Terrier, Chen, and Sutter, PNAS 2021
  - ed tech records and administrative data

    Chen, Ertac, Evgeniou, Nadaf, Miao, Yilmax, Nature Education 2024

AI Education as State Capacity: Experimental Evidence from Pakistan, Mehmood, Naseer, and Chen, J Development Econ R&R

The Effects of Entrepreneurship Education, Von Graevenitz, Harhoff, Weber, J Econ Behavioral Org 2010

Fostering Innovation

- legal doctrine, security of property rights, and trust in the law
  - how these can stimulate investment, reliance, growth
  - these are standard ideas that also work in ideas creation
- **Fiscal Policy**:
  - summary of the literature finds that a 10 percent fall in the tax price of R&D generates a 10 percent increase in R&D in the long-run

    Bloom, Van Reenen, and Williams, Journal of Econ Perspectives 2019
  - alternatives, e.g., prizes, contests, patent buyouts, open access
- **Social Policy**: individual determinants of innovation

  Bell, Chetty, Jaravel, Petkova, Van Reenan, Quarterly J Econ 2019
  - intrinsic motivation (social emotional learning measurements)

    Terrier, Chen, and Sutter, PNAS 2021
  - ed tech records and administrative data

    Chen, Ertac, Evgeniou, Nadaf, Miao, Yilmax, Nature Education 2024

AI Education as State Capacity: Experimental Evidence from Pakistan, Mehmood, Naseer, and Chen, J Development Econ R&R

The Effects of Entrepreneurship Education, Von Graevenitz, Harhoff, Weber, J Econ Behavioral Org 2010

Recap

- **Economics + Law => AI**
  - Kantian LLMs
- **AI + Economics => Law**
  - Nationwide RCT ↓case duration by 20% and ↑wages by 8%
- **AI + Law => Economics**
  - Brain tech & cognitive economics

Chen and Schonger, Social Preferences or Sacred Values? Theory and Evidence of Deontological Motivations. Science Advances 2022

Innovation and Competition: AI, Law, and Economics

- **AI Safety** *Economics + Law => AI*
  - normative commitments
  - oTree/Ed: behavioral research platform
  - legal frameworks for the digital economy [Drexl]
- **AMICUS** (Analytical Metrics for Informed Court Understanding & Strategy) *AI + Economics => Law*
  - causal impacts of decisions
  - causal impacts of texts
  - causal impacts of institutions [Hilty]
- **AI as Motor of Innovation** [Harhoff] *AI + Law => Economics*
  - self-GPT
  - mind-body economics
  - foundation fNIRS models & decision-making

Chen and Reinhart, Markets and Morality: How Markets Shape Our (Dis)regard for Others. J of Law, Econ, Org 2024

Innovation and Competition: AI, Law, and Economics

- **AI Safety** *Economics + Law => AI*
  - normative commitments
  - oTree/Ed: behavioral research platform
  - legal frameworks for the digital economy [Drexl]
- **AMICUS** (Analytical Metrics for Informed Court Understanding & Strategy) *AI + Economics => Law*
  - causal impacts of decisions
  - causal impacts of texts
  - causal impacts of institutions [Hilty]
- **AI as Motor of Innovation** [Harhoff] *AI + Law => Economics*
  - self-GPT
  - mind-body economics
  - foundation fNIRS models & decision-making

Chen and Reinhart, Markets and Morality: How Markets Shape Our (Dis)regard for Others. J of Law, Econ, Org 2024

Innovation and Competition: AI, Law, and Economics

- **AI Safety** *Economics + Law => AI*
  - normative commitments
  - oTree/Ed: behavioral research platform
  - legal frameworks for the digital economy [Drexl]
- **AMICUS** (Analytical Metrics for Informed Court Understanding & Strategy) *AI + Economics => Law*
  - causal impacts of decisions
  - causal impacts of texts
  - causal impacts of institutions [Hilty]
- **AI as Motor of Innovation** [Harhoff] *AI + Law => Economics*
  - self-GPT
  - mind-body economics
  - foundation fNIRS models & decision-making

Chen and Reinhart, Markets and Morality: How Markets Shape Our (Dis)regard for Others. J of Law, Econ, Org 2024

Innovation and Competition: AI, Law, and Economics

- **AI Safety** *Economics + Law => AI*
  - normative commitments
  - oTree/Ed: behavioral research platform
  - legal frameworks for the digital economy [Drexl]
- **AMICUS** (Analytical Metrics for Informed Court Understanding & Strategy) *AI + Economics => Law*
  - causal impacts of decisions
  - causal impacts of texts
  - causal impacts of institutions [Hilty]
- **AI as Motor of Innovation** [Harhoff] *AI + Law => Economics*
  - self-GPT
  - mind-body economics
  - foundation fNIRS models & decision-making

Chen and Reinhart, Markets and Morality: How Markets Shape Our (Dis)regard for Others. J of Law, Econ, Org 2024