Optimal Taxation in the Age of AI Uncertainty

Anastasios Karantounias[1] Radek Paluszynski[2] PC Yu[3]

[1]University of Surrey

[2]University of Houston

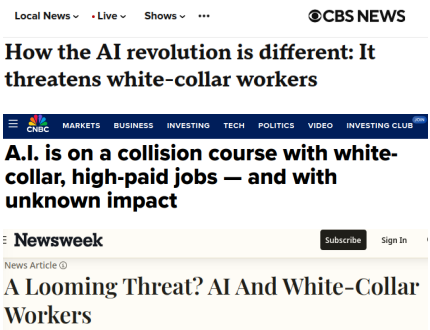
[3]UNSW

Conference on Research on Economic Theory and Econometrics

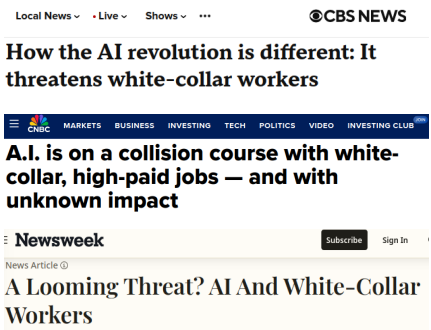
July 2026

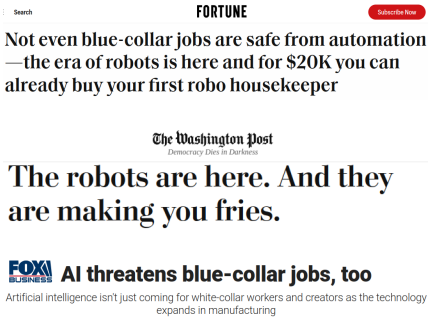
Motivation: substantial uncertainty about new technologies

Motivation: substantial uncertainty about new technologies



Motivation: substantial uncertainty about new technologies





This Paper

- How should we design policies in the face of *unknown* risks/uncertainty?

This Paper

- How should we design policies in the face of *unknown* risks/uncertainty?
- Ambiguity of workers about an *aggregate* technological shock.

This Paper

- How should we design policies in the face of *unknown* risks/uncertainty?
- Ambiguity of workers about an *aggregate* technological shock.
- *Optimal* Mirrlees taxation with *ambiguity-averse* workers.

This Paper

- How should we design policies in the face of *unknown* risks/uncertainty?
- Ambiguity of workers about an *aggregate* technological shock.
- *Optimal* Mirrlees taxation with *ambiguity-averse* workers.
- Questions:
  - How should we tax income in the face of ambiguity?
  - Are there incentives for distorting savings?

Preview of results

- Ambiguity: theory of endogenous and heterogenous *worst-case* beliefs.

Preview of results

- Ambiguity: theory of endogenous and heterogenous *worst-case* beliefs.
  - Agents assign *high* probability on *low* utility events.

Preview of results

- Ambiguity: theory of endogenous and heterogenous *worst-case* beliefs.
  - Agents assign *high* probability on *low* utility events.
  - Worst-case beliefs *differ* across the skill distribution.

Preview of results

- Ambiguity: theory of endogenous and heterogenous *worst-case* beliefs.
  - Agents assign *high* probability on *low* utility events.
  - Worst-case beliefs *differ* across the skill distribution.
  - Optimal policy: Planner *manipulates* beliefs to improve efficiency and welfare.
  - $\Rightarrow$ *Novel* incentives to distort *intertemporal* and *intratemporal* margins.

Preview of results

- Ambiguity: theory of endogenous and heterogenous *worst-case* beliefs.
  - Agents assign *high* probability on *low* utility events.
  - Worst-case beliefs *differ* across the skill distribution.
  - Optimal policy: Planner *manipulates* beliefs to improve efficiency and welfare.
  - $\Rightarrow$ *Novel* incentives to distort *intertemporal* and *intratemporal* margins.
- Quantitative application on AI uncertainty `(preliminary)`:
  - Ambiguity about the *substitutability* of skill and AI.

Preview of results

- Ambiguity: theory of endogenous and heterogenous *worst-case* beliefs.
  - Agents assign *high* probability on *low* utility events.
  - Worst-case beliefs *differ* across the skill distribution.
  - Optimal policy: Planner *manipulates* beliefs to improve efficiency and welfare.
  - $\Rightarrow$ *Novel* incentives to distort *intertemporal* and *intratemporal* margins.
- Quantitative application on AI uncertainty `(preliminary)`:
  - Ambiguity about the *substitutability* of skill and AI.
  - Use *survey* evidence to discipline belief dispersion.

Preview of results

- Ambiguity: theory of endogenous and heterogenous *worst-case* beliefs.
  - Agents assign *high* probability on *low* utility events.
  - Worst-case beliefs *differ* across the skill distribution.
  - Optimal policy: Planner *manipulates* beliefs to improve efficiency and welfare.
  - $\Rightarrow$ *Novel* incentives to distort *intertemporal* and *intratemporal* margins.
- Quantitative application on AI uncertainty `(preliminary)`:
  - Ambiguity about the *substitutability* of skill and AI.
  - Use *survey* evidence to discipline belief dispersion.
  - Findings:
    - Near flat savings *subsidy* around 13%.
    - The labor tax becomes *more* progressive.
    - Current policies: "pessimism" *negatively* correlated with income.
    - Optimal system: "pessimism" *positively* correlated with income.

Model

Economy

- Two periods: $t = 0, 1$ with linear storage technology (return R).
- Permanent skill: $\theta \in \Theta = \left[\underline{\theta}, \overline{\theta}\right]$ with cdf $F(\theta)$.

Economy

- Two periods: $t = 0, 1$ with linear storage technology (return R).
- Permanent skill: $\theta \in \Theta = \left[\underline{\theta}, \overline{\theta}\right]$ with cdf $F(\theta)$.
- Technology:
  - At $t = 0 : y = \theta\ell$, θ and ℓ *not* observable.

Economy

- Two periods: $t = 0, 1$ with linear storage technology (return R).
- Permanent skill: $\theta \in \Theta = \left[\underline{\theta}, \overline{\theta}\right]$ with cdf $F(\theta)$.
- Technology:
  - At $t = 0 : y = \theta\ell$, θ and ℓ *not* observable.
  - At $t = 1 : y = \omega(\theta, \xi)\,\ell$
  - *Productivity* (wage) $\omega(\theta, \xi)$:
    - Aggregate technology shock ξ.
    - $\frac{\partial \omega}{\partial \theta} > 0$, $\frac{\partial \omega}{\partial \xi} > 0$.
    - Ambiguity about ξ.

Economy

- Two periods: $t = 0, 1$ with linear storage technology (return R).
- Permanent skill: $\theta \in \Theta = \left[\underline{\theta}, \overline{\theta}\right]$ with cdf $F(\theta)$.
- Technology:
  - At $t = 0 : y = \theta\ell$, θ and ℓ *not* observable.
  - At $t = 1 : y = \omega(\theta, \xi)\,\ell$
  - *Productivity* (wage) $\omega(\theta, \xi)$:
    - Aggregate technology shock ξ.
    - $\frac{\partial \omega}{\partial \theta} > 0$, $\frac{\partial \omega}{\partial \xi} > 0$.
    - Ambiguity about ξ.
  - Resource constraint:

$$\int_{\Theta} \left\{ c_0(\theta) + \frac{c_1(\theta, \xi_i)}{R} \right\} dF(\theta) \leq \int_{\Theta} \left\{ y_0(\theta) + \frac{y_1(\theta, \xi_i)}{R} \right\} dF(\theta), i = 1, ..., N$$

Ambiguity aversion I

- Expected utility: DM has a unique probability model π that he completely trusts.

$$U(c) = \int u(c) d\pi$$

Ambiguity aversion I

- Expected utility: DM has a unique probability model π that he completely trusts.

$$U(c) = \int u(c) d\pi$$

- Doubts about the model/ambiguity: set of probability models.

Ambiguity aversion I

- Expected utility: DM has a unique probability model π that he completely trusts.

$$U(c) = \int u(c) d\pi$$

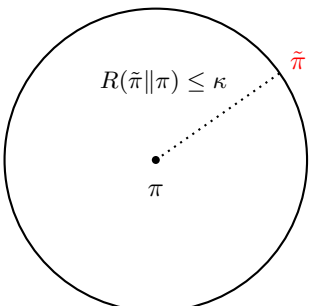
- Doubts about the model/ambiguity: set of probability models.
  - DM has a *reference* model π that he does not fully trust.

Ambiguity aversion I

- Expected utility: DM has a unique probability model π that he completely trusts.

$$U(c) = \int u(c) d\pi$$

- Doubts about the model/ambiguity: set of probability models.
  - DM has a *reference* model π that he does not fully trust.
  - DM considers *alternative* $\tilde{\pi}$ "close" to π.

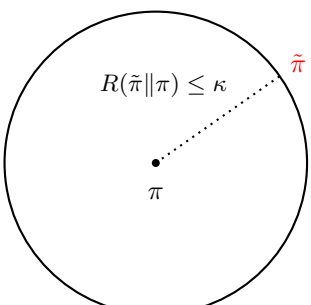


Ambiguity aversion I

- Expected utility: DM has a unique probability model π that he completely trusts.

$$U(c) = \int u(c) d\pi$$

- Doubts about the model/ambiguity: set of probability models.
  - DM has a *reference* model π that he does not fully trust.
  - DM considers *alternative* $\tilde{\pi}$ "close" to π.



- *Multiplier preferences* of Hansen and Sargent (2001)

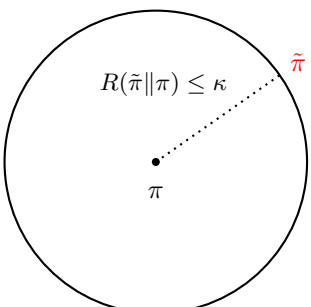
$$U(c) = \min_{\tilde{\pi}} \int u(c) d\tilde{\pi} + \chi \cdot R(\tilde{\pi}\|\pi)$$

Ambiguity aversion I

- Expected utility: DM has a unique probability model π that he completely trusts.

$$U(c) = \int u(c) d\pi$$

- Doubts about the model/ambiguity: set of probability models.
  - DM has a *reference* model π that he does not fully trust.
  - DM considers *alternative* $\tilde{\pi}$ "close" to π.



- *Multiplier preferences* of Hansen and Sargent (2001)

$$U(c) = \min_{\tilde{\pi}} \int u(c) d\tilde{\pi} + \chi \cdot R(\tilde{\pi}\|\pi)$$

- $\chi > 0$: penalty parameter that captures distrust of the model
- $\chi \to \infty \Rightarrow \tilde{\pi} = \pi$, *full* confidence in π (EU).

Ambiguity aversion II

- Each worker with skill θ.
  - Shares the *same* reference π about the shock ξ.
  - Alternative $\tilde{\pi}(\theta) \Rightarrow$ likelihood ratio $m_i(\theta) \equiv \frac{\tilde{\pi}_i(\theta)}{\pi_i} \geq 0, \quad \sum_i \pi_i m_i(\theta) = 1.$

Ambiguity aversion II

- Each worker with skill θ.
  - Shares the *same* reference π about the shock ξ.
  - Alternative $\tilde{\pi}(\theta) \Rightarrow$ likelihood ratio $m_i(\theta) \equiv \frac{\tilde{\pi}_i(\theta)}{\pi_i} \geq 0, \quad \sum_i \pi_i m_i(\theta) = 1.$
- Preferences

$$u(c_0(\theta)) - h\left(\frac{y_0(\theta)}{\theta}\right) + \beta \min_{m_i(\theta)} \quad \Big\{ \sum_i \pi_i m_i(\theta) \left[u(c_i(\theta, \xi_i)) - h\left(\tfrac{y_1(\theta,\xi_i)}{\omega(\theta,\xi_i)}\right) \right] + \chi \cdot \underbrace{\sum_i \pi_i m_i(\theta) \ln m_i(\theta)}_{\text{relative entropy } R(\tilde{\pi}(\theta) \| \pi)} \Big\}$$

Worst-case beliefs

- Worst-case $m_i(\theta)$:

$$m_i(\theta) = \frac{\exp\big(\sigma\big[u(c_i(\theta,\xi_i)) - h(\frac{y_1(\theta,\xi_i)}{\omega(\theta,\xi_i)})\big]\big)}{\sum_i \pi_i \exp\big(\sigma\big[u(c_i(\theta,\xi_i)) - h(\frac{y_1(\theta,\xi_i)}{\omega(\theta,\xi_i)})\big]} \quad \forall i$$

- $\sigma \equiv -1/\chi < 0$. If $\sigma = 0$ (*full* confidence) $\Rightarrow m_i(\theta) = 1 \forall i, \theta$.

Worst-case beliefs

- Worst-case $m_i(\theta)$:

$$m_i(\theta) = \frac{\exp\big(\sigma\big[u(c_i(\theta,\xi_i)) - h(\frac{y_1(\theta,\xi_i)}{\omega(\theta,\xi_i)})\big]\big)}{\sum_i \pi_i \exp\big(\sigma\big[u(c_i(\theta,\xi_i)) - h(\frac{y_1(\theta,\xi_i)}{\omega(\theta,\xi_i)})\big]} \quad \forall i$$

- $\sigma \equiv -1/\chi < 0$. If $\sigma = 0$ (*full* confidence) $\Rightarrow m_i(\theta) = 1 \forall i, \theta$.
- Endogeneity: Agents assign *high* probability on *low* utility events.
- Heterogeneity: worst-case beliefs *differ* across types θ.

Worst-case beliefs

- Worst-case $m_i(\theta)$:

$$m_i(\theta) = \frac{\exp\big(\sigma\big[u(c_i(\theta,\xi_i)) - h(\frac{y_1(\theta,\xi_i)}{\omega(\theta,\xi_i)})\big]\big)}{\sum_i \pi_i \exp\big(\sigma\big[u(c_i(\theta,\xi_i)) - h(\frac{y_1(\theta,\xi_i)}{\omega(\theta,\xi_i)})\big]}, \ \forall i$$

- $\sigma \equiv -1/\chi < 0$. If $\sigma = 0$ (*full* confidence) $\Rightarrow m_i(\theta) = 1 \forall i, \theta$.
- Endogeneity: Agents assign *high* probability on *low* utility events.
- Heterogeneity: worst-case beliefs *differ* across types θ.
- Indirect utility of type θ:

$$U_0(\theta) = u(c_0(\theta)) - h(\frac{y_0(\theta)}{\theta}) + \frac{\beta}{\sigma} \ln \sum_i \pi_i \exp\left(\sigma\Big[u(c_i(\theta,\xi_i)) - h(\frac{y_1(\theta,\xi_i)}{\omega(\theta,\xi_i)}\Big)\Big]\right)$$

- Connection with *risk-sensitive* preferences.

Utilitarian Planner's Problem

Let $\beta = 1/R$. Choose $\{c_0(\theta), c_1(\theta, \xi_i), y_0(\theta), y_1(\theta, \xi_i)\}_{\forall\theta,\xi_i}$ to max

$$\int_{\Theta} U_0(\theta)\, dF(\theta)$$

subject to

1. *Incentive compatibility* constraints

$$U_0(\theta, \theta) \geq U_0(\theta, \hat{\theta}), \forall\theta, \hat{\theta}$$

2. The resource constraint $\forall\xi_i$.

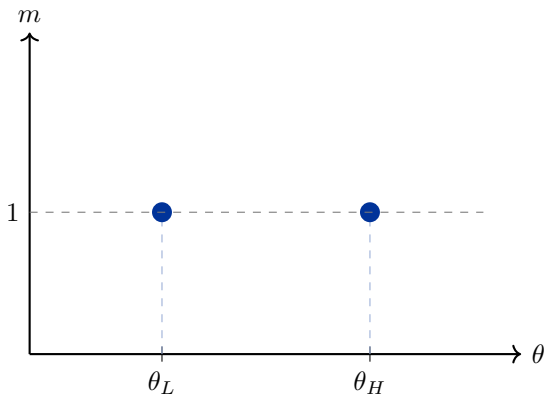
- State *non-separabilities* through the *endogenous* beliefs of the agents.
- Optimal policy: *manipulate* the beliefs of workers to relax ICs.

Full-information benchmark

Optimal Policies

How Beliefs Shift When Agents Misreport

Fix ξ_i.

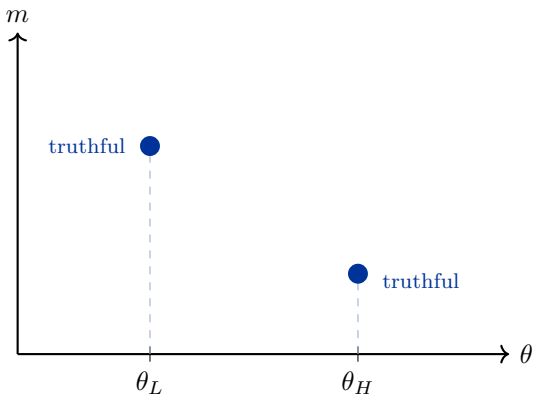


No ambiguity aversion:

Beliefs are the same for all

How Beliefs Shift When Agents Misreport

Fix ξ_i.

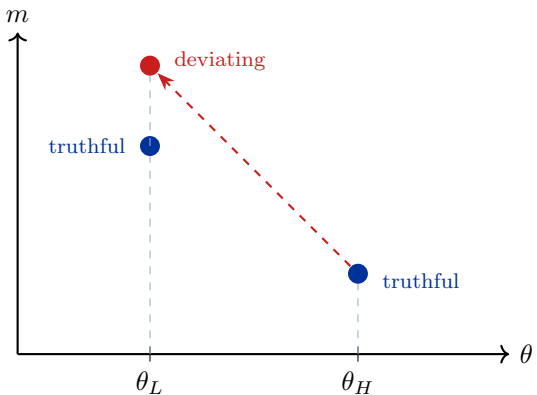


With ambiguity aversion:

Beliefs *differ* across agents

How Beliefs Shift When Agents Misreport

Fix ξ_i.



Case 1

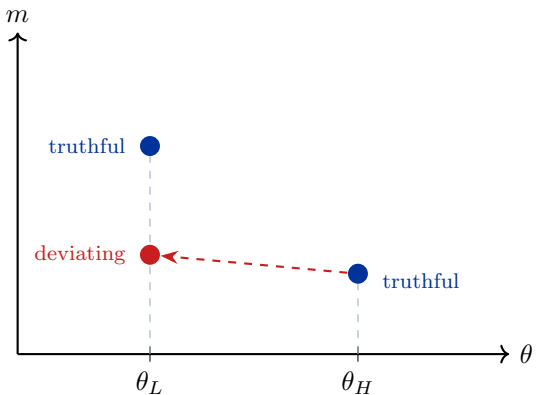
Deviating agent has *higher* MU than truthful agent

$\Rightarrow$ *Punish* the deviating agent:

- *Tax* savings
- *Increase* labor wedge

How Beliefs Shift When Agents Misreport

Fix ξ_i.



Case 2

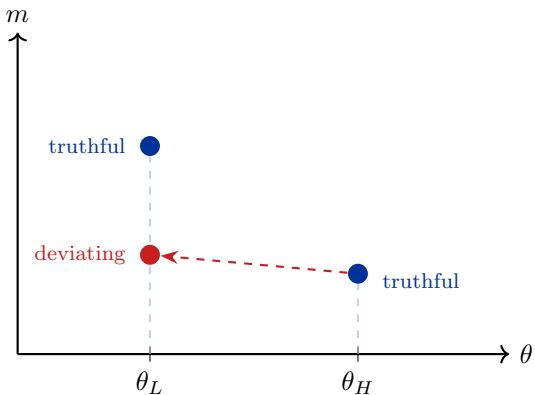
Deviating agent has *smaller* MU than the truthful agent

$\Rightarrow$ *Reward* truthful agent:

- *Subsidize* savings
- *Reduce* labor wedge

How Beliefs Shift When Agents Misreport

Fix ξ_i.



Case 2

Deviating agent has *smaller* MU than the truthful agent

$\Rightarrow$ *Reward* truthful agent:

- *Subsidize* savings
- *Reduce* labor wedge

Key: Tax (subsidize) what the deviating agent likes (dislikes)

Belief Wedge Between Honest and Dishonest Agents

Let $m_i(\theta, \hat{\theta})$ the likelihood ratio of of θ who reports $\hat{\theta}$.

$$\frac{\partial m_i(\theta, \hat{\theta})}{\partial \theta}_{|\hat{\theta}=\theta} \cdot \frac{\theta}{m_i(\theta)} = \sigma \cdot \eta_i(\theta)$$

where

$$\eta_i(\theta) \equiv \underbrace{h'(\ell_1(\theta, \xi_i))\ell_1(\theta, \xi_i)\epsilon^{\omega}_{\theta}(\theta, \xi_i)}_{\text{MB of misreporting at } \xi_i} - \underbrace{\tilde{E}_{\theta} h'(\ell_1(\theta, \xi))\ell_1(\theta, \xi)\epsilon^{\omega}_{\theta}(\theta, \xi)}_{\text{Average MB of misreporting}}$$

with $\sum_i p_i m_i(\theta)\eta_i(\theta) = 0 \Rightarrow$ *innovation* in MB of misreporting under worst-case beliefs.

Belief Wedge Between Honest and Dishonest Agents

Let $m_i(\theta, \hat{\theta})$ the likelihood ratio of of θ who reports $\hat{\theta}$.

$$\frac{\partial m_i(\theta, \hat{\theta})}{\partial \theta}_{|\hat{\theta}=\theta} \cdot \frac{\theta}{m_i(\theta)} = \sigma \cdot \eta_i(\theta)$$

where

$$\eta_i(\theta) \equiv \underbrace{h'(\ell_1(\theta, \xi_i))\ell_1(\theta, \xi_i)\epsilon^{\omega}_{\theta}(\theta, \xi_i)}_{\text{MB of misreporting at } \xi_i} - \underbrace{\tilde{E}_{\theta} h'(\ell_1(\theta, \xi))\ell_1(\theta, \xi)\epsilon^{\omega}_{\theta}(\theta, \xi)}_{\text{Average MB of misreporting}}$$

with $\sum_i p_i m_i(\theta)\eta_i(\theta) = 0 \Rightarrow$ *innovation* in MB of misreporting under worst-case beliefs.

- Difference between misreporting and truthful

$$m_i(\theta + \Delta\theta, \theta) - m_i(\theta) \simeq \frac{\partial m_i(\theta, \hat{\theta})}{\partial \theta}_{|\hat{\theta}=\theta} \cdot \Delta\theta \propto \sigma\eta_i(\theta)$$

- If $\eta_i(\theta) > 0 \Rightarrow m_i(\theta + \Delta\theta, \theta) < m_i(\theta)$ (Case 2)
- If $\eta_i(\theta) < 0 \Rightarrow m_i(\theta + \Delta\theta, \theta) > m_i(\theta)$ (Case 1)

Novel savings wedge

- *Ambiguity-tilted* Euler equation:

$$u'\left(c_0\left(\theta\right)\right) = \sum_{i=0}^{S} \pi_i m_i(\theta) u'\left(c_1\left(\theta, \xi_i\right)\right) \cdot \Gamma\left(\theta, \xi_i\right)$$

- If $\chi \to \infty \Rightarrow \Gamma(\theta, \xi_i) = 1 \forall \theta, \xi_i \Rightarrow$ *no* wedge!

Novel savings wedge

- *Ambiguity-tilted* Euler equation:

$$u'\left(c_0\left(\theta\right)\right) = \sum_{i=0}^{S} \pi_i m_i(\theta) u'\left(c_1\left(\theta, \xi_i\right)\right) \cdot \Gamma\left(\theta, \xi_i\right)$$

- If $\chi \to \infty \Rightarrow \Gamma(\theta, \xi_i) = 1 \forall \theta, \xi_i \Rightarrow$ *no* wedge!
- Type and state-contingent wedge: $\tau^s(\theta, \xi_i) \equiv 1 - \Gamma(\theta, \xi_i)$

Novel savings wedge

- *Ambiguity-tilted* Euler equation:

$$u'\left(c_0\left(\theta\right)\right) = \sum_{i=0}^{S} \pi_i m_i(\theta) u'\left(c_1\left(\theta, \xi_i\right)\right) \cdot \Gamma\left(\theta, \xi_i\right)$$

- If $\chi \to \infty \Rightarrow \Gamma(\theta, \xi_i) = 1 \forall \theta, \xi_i \Rightarrow$ *no* wedge!
- Type and state-contingent wedge: $\tau^s(\theta, \xi_i) \equiv 1 - \Gamma(\theta, \xi_i)$
  - If $\eta_i(\theta) > 0 \Rightarrow \tau^s(\theta, \xi_i) < 0$ (subsidy on savings).
  - If $\eta_i(\theta) < 0 \Rightarrow \tau^s(\theta, \xi_i) > 0$ (tax on savings).

Novel savings wedge

- *Ambiguity-tilted* Euler equation:

$$u'\left(c_0\left(\theta\right)\right) = \sum_{i=0}^{S} \pi_i m_i(\theta) u'\left(c_1\left(\theta, \xi_i\right)\right) \cdot \Gamma\left(\theta, \xi_i\right)$$

- If $\chi \to \infty \Rightarrow \Gamma(\theta, \xi_i) = 1 \forall \theta, \xi_i \Rightarrow$ *no* wedge!
- Type and state-contingent wedge: $\tau^s(\theta, \xi_i) \equiv 1 - \Gamma(\theta, \xi_i)$
  - If $\eta_i(\theta) > 0 \Rightarrow \tau^s(\theta, \xi_i) < 0$ (subsidy on savings).
  - If $\eta_i(\theta) < 0 \Rightarrow \tau^s(\theta, \xi_i) > 0$ (tax on savings).
- *Distort* savings to *relax* ICs $\Rightarrow$ punish *lying*/ reward *truth-telling*.
  - $\eta_i(\theta) > 0 \Rightarrow m_i(\text{truthful}) \uparrow \Rightarrow$ *subsidize* savings $\Rightarrow U(\text{truthful}) \uparrow$
  - $\eta_i(\theta) < 0 \Rightarrow m_i(\text{lying}) \uparrow \Rightarrow$ *tax* savings $\Rightarrow U(\text{lying}) \downarrow$
- Savings wedge: $\tau^s(\theta) \equiv \sum_i \pi_i \nu_i(\theta) \tau^s(\theta, \xi_i)$, where $\nu_i(\theta)$: *ambiguity* and *risk-adjusted* change of measure. ⏩ Labor wedge in $t = 0$ ⏩ Labor wedge in $t = 1$

Quantitative application

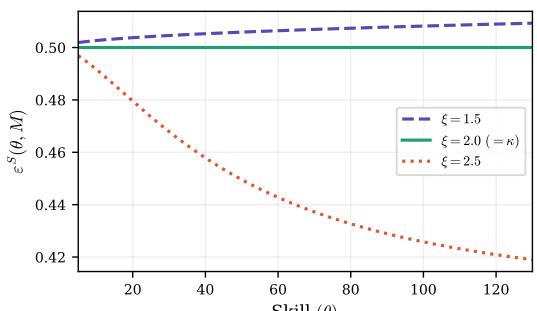
AI capital and Technology Shocks

- Productivity: $\omega = \omega(M, \theta, \xi)$ with *"AI capital"* M with price ϕ.
  - M mixes with θ and ξ, $\frac{\partial \omega}{\partial M} > 0$

AI capital and Technology Shocks

- Productivity: $\omega = \omega(M, \theta, \xi)$ with *"AI capital"* M with price ϕ.
  - M mixes with θ and ξ, $\frac{\partial \omega}{\partial M} > 0$
  - ξ affects *substitutability* between (θ, M).

$$\omega\left(M, \theta, \xi\right) = \left[\alpha \theta^{1-\xi} + (1-\alpha) M^{1-\kappa}\right]^{\frac{1}{1-\kappa}}$$



- Set $\kappa = 2$ (high Hicksian complementarity)
- $\xi = \kappa$: CES benchmark
- $\xi > \kappa$: complementarity *increases* with skill
- $\xi < \kappa$: complementarity *decreases* with skill

Calibration

- "Current policy" economy with HSV(2017) policies, CRRA/constant Frisch.
- Calibrate $F(\theta)$ to match earnings. ⏩ calibration
- "Pessimism":

$$\text{Pessimism}_j \equiv 1 - \frac{\tilde{\mathbb{E}}(y_{1,j})}{\mathbb{E}_\pi(y_{1,j})}$$

- Calibrate $\{\alpha, \pi, \chi\}$ to match pessimism in the *current policy* economy with ⏩ USC survey evidence on the impact of AI.

Calibration

- "Current policy" economy with HSV(2017) policies, CRRA/constant Frisch.
- Calibrate $F(\theta)$ to match earnings. ⏩ calibration
- "Pessimism":

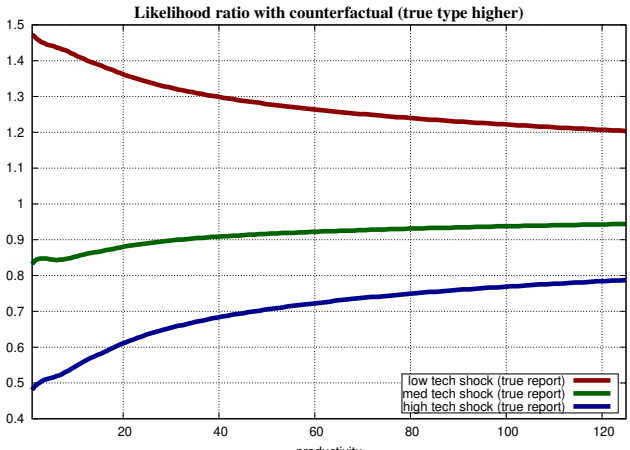
$$\text{Pessimism}_j \equiv 1 - \frac{\tilde{\mathbb{E}}(y_{1,j})}{\mathbb{E}_\pi(y_{1,j})}$$

- Calibrate $\{\alpha, \pi, \chi\}$ to match pessimism in the *current policy* economy with ⏩ USC survey evidence on the impact of AI.

Param.	Meaning	Value	Moment	Current policy	Survey
χ	Ambiguity aversion	0.15	Mean *pessimism*	0.23	0.19
α	Share of θ	0.14	St. dev. *pessimism*	0.19	0.25
σ_ξ	Scale parameter	0.71	Pessimism-log(inc) coeff.	-0.07	-0.07
β_ξ	Shape parameter	3.94	$\text{Corr}(y_0, \frac{\mathbb{E}_p(y_1)}{y_0})$	0.00	-
μ_ξ	Location parameter	0.71	$\mathbb{E}_p(\xi) - \kappa$	0.00	-

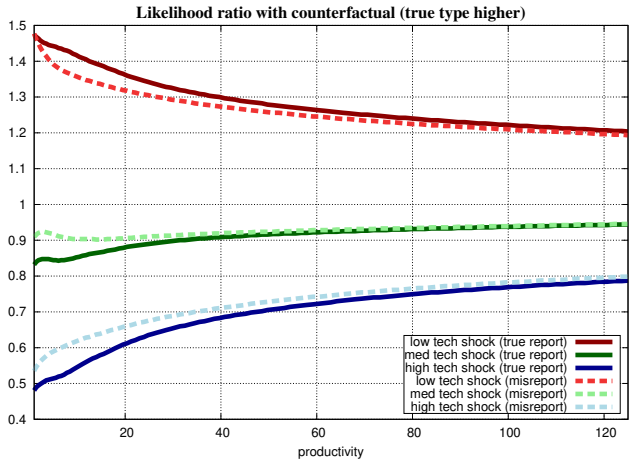
⏩ Distribution of ξ

Endogenous Subjective Beliefs



- *High* m_i for *low* ξ_i (low complementarity)
  $\Rightarrow$ "bad" shock.
- *Low* m_i for *high* ξ_i (high complementarity)
  $\Rightarrow$ "good" shock.

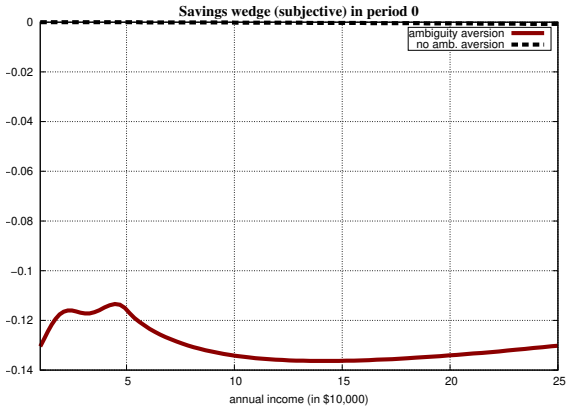
Endogenous Subjective Beliefs



- Deviating agent places more weight on *high* shocks than truthful
  $\Rightarrow$ *tax*
- Deviating agent places less weight on *low* shocks than truthful
  $\Rightarrow$ *subsidize*

Optimal distortions

(a) Labor wedge



(b) Savings wedge

- Optimal labor wedge $\Longrightarrow$ more progressive with amb. aversion Labor wedge t=1
- Overall savings wedge is *negative*: around 13% subsidy

Belief manipulation

Pessimism

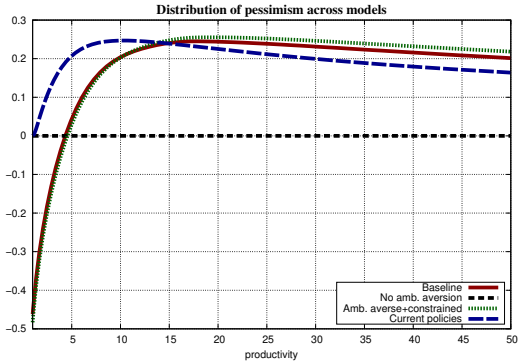
	Mean	Std Dev	Pessimism-income reg coeff
"Current policy" model	0.23	0.19	-0.07
Baseline ambiguity aversion	0.04	0.22	0.28
No ambiguity aversion	0.00	0.00	-

- Mean "pessimism" reduced in optimal system (variance the same).
- "Pessimism" positively correlated with income.

Belief manipulation

Pessimism

	Mean	Std Dev	Pessimism-income reg coeff
"Current policy" model	0.23	0.19	-0.07
Baseline ambiguity aversion	0.04	0.22	0.28
No ambiguity aversion	0.00	0.00	-



Concluding remarks

- Characterize *Mirrlees* policies with ambiguity aversion
  - Ambiguity aversion opens door for belief manipulation
  - Novel distortions at the *intertemporal* and *intratemporal* margin.
- Quantitative effects of belief manipulation are *non-negligible*
  - 13% savings subsidy.
  - More progressive income taxes.
  - Reversion of "pessimism" pattern.

THANK YOU!

Related Literature

Ramsey taxation with multiplier preferences:

- Karantounias (2013, 2023), Ferriere & Karantounias (2019)

Risk sharing with multiplier preferences:

- Anderson (2005), Bhandari (2015), Miao & Rivera (2016)

Mirrlees taxation with recursive preferences:

- Farhi & Werning (2008), Kapicka (2024)

Mirrlees taxation with planner ambiguity:

- Bhandari, Borovicka & Yao (2025), Vairo (2025)

Mirrlees taxation with non-separable preferences:

- Mirrlees (1971), Saez (2001), Golosov, Kocherlakota & Tsyvinski (2006), Farhi & Werning (2013), Golosov, Troshkin & Tsyvinski (2016), Hellwig (2021)

Back

USC data

- Understanding America Study (UAS): nationally representative internet-based panel
- Survey by Center for Economic and Social Research (CESR) at University of Southern California
- Artificial Intelligence wave 3: fielded February-May 2025
- About 10,000 respondents, paid $8 to participate

How will AI affect probability of losing a job?

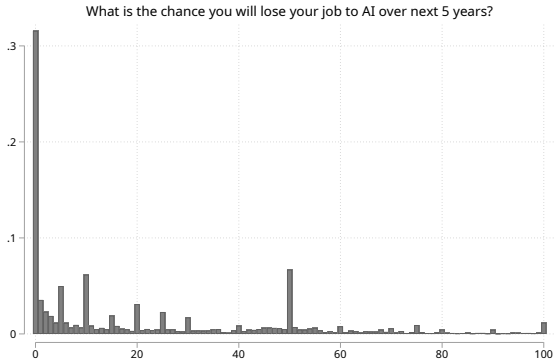


Figure A: Raw responses

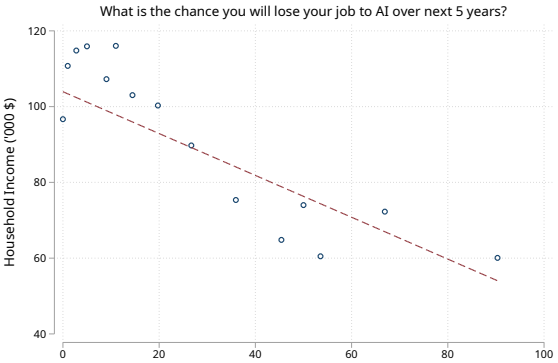


Figure B: Responses by income

- Rare numerical response: use in calibration (*"pessimism"*)
- Correlation between job loss probability and income -0.9
- Alternative source: Pew Research Center survey Oct 2024 - similar!

Back

How will AI affect the level of pay?

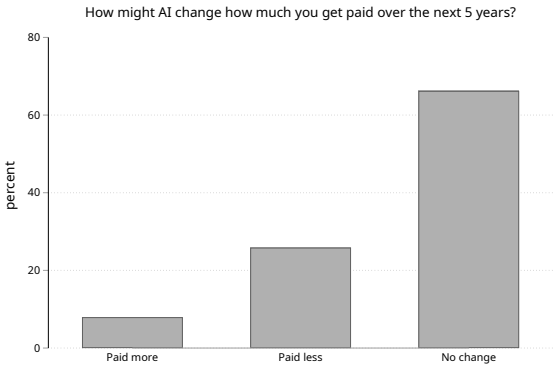


Figure A: Raw responses

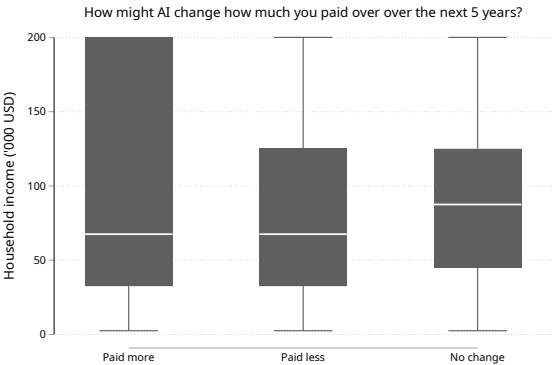


Figure B: Responses by income

Similar response patterns: How much will AI affect the number of jobs?

How much will AI affect how much you accomplish at work?

How much will AI affect how much you enjoy your job?

How will AI affect the number of jobs?

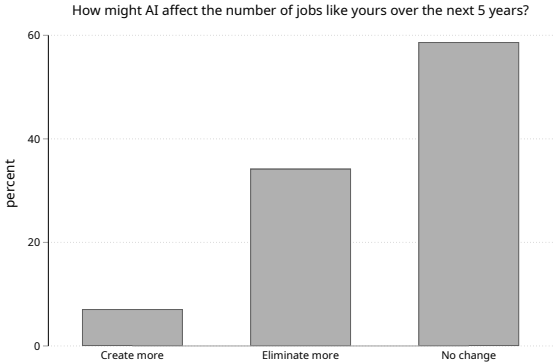


Figure A: Raw responses

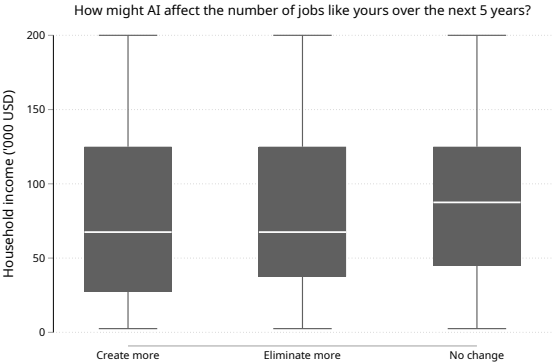


Figure B: Responses by income

Back

How will AI affect what you can accomplish at work?

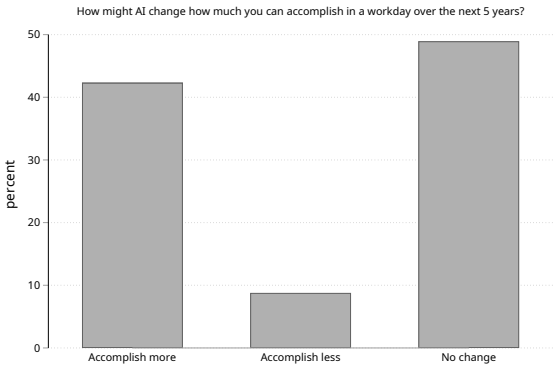


Figure A: Raw responses

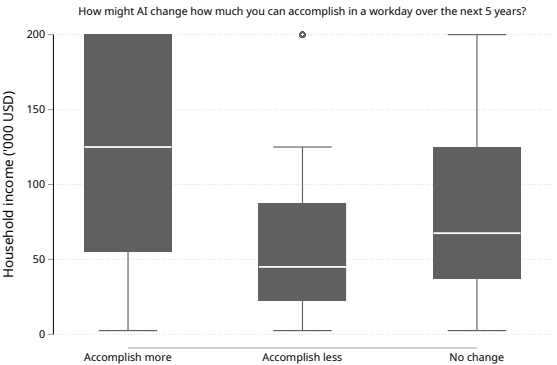


Figure B: Responses by income

Back

How will AI affect how you enjoy your job?

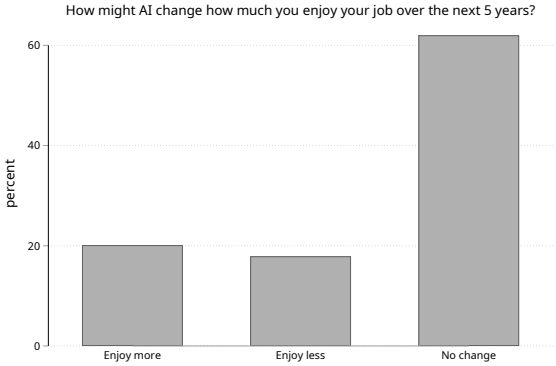


Figure A: Raw responses

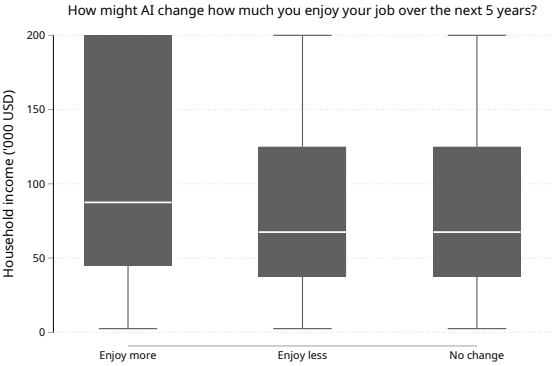


Figure B: Responses by income

Back

What fraction of job tasks could be automated?

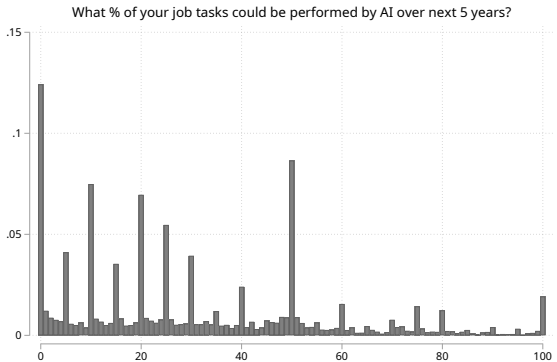


Figure A: Raw responses

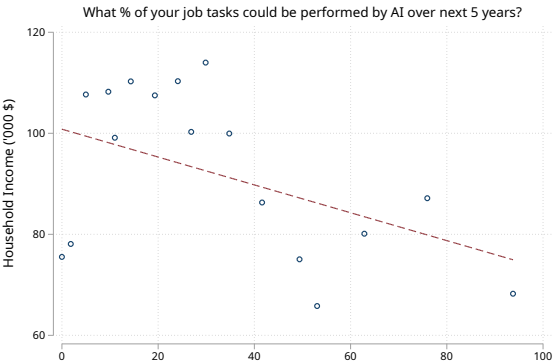


Figure B: Responses by income

Back

Full-Information Benchmark

1. Belief-adjusted Euler equation:

$$u'(c_0(\theta)) = \sum_{i=0}^{S} \pi_i \, m_i(\theta) u'(c_1(\theta, \xi_i))$$

2. No wedges in labor supply at $t = 0$ and all ξ_i at $t = 1$, $\forall \theta$.

3. Full insurance at $t = 0$: $c_0(\theta) = c_0(\theta') = c_0 \ \forall \theta, \theta'$.

4. Belief-adjusted insurance for all ξ_i and $\theta' \in \Theta$ at $t = 1$:

$$\frac{m_i(\theta) u'(c_1(\theta, \xi_i))}{m_i(\theta') u'(c_1(\theta', \xi_i))} = 1 \quad \forall \xi_i, \theta, \theta'$$

Back

Income Tax at $t = 0$ (Mirrlees-Saez-Diamond Formula)

$$\frac{\tau_0^w(\theta)}{1-\tau_0^w(\theta)} = A_0(\theta) B_0(\theta) C_0(\theta)$$

where

$$A_0(\theta) = \frac{1-F(\theta)}{\theta f(\theta)}$$

$$B_0(\theta) = 1 + \frac{1}{\gamma_0(\theta)}$$

$$C_0(\theta) = \int_\theta^{\overline{\theta}} \frac{u'(c_0(\theta))}{u'(c_0(x))} \left[1 - \frac{u'(c_0(x))}{\sum_{i=0}^S \mu(\xi_i)}\right] \frac{f(x)}{1-F(\theta)} dx$$

- Frisch elasticity $\gamma_0(\theta) = h'\left(\frac{y_0(\theta)}{\theta}\right) / \frac{y_0(\theta)}{\theta} h''\left(\frac{y_0(\theta)}{\theta}\right)$
- $\mu(\xi_i)$ is the multiplier on the resource constraint

Back

Labor wedge at $t = 1$ for ξ_i: Modified ABC formula

$$\frac{\tau_1^w(\theta, \xi_i)}{1 - \tau_1^w(\theta, \xi_i)} = \underbrace{\frac{1 - F(\theta)}{\theta f(\theta)} \cdot \epsilon_\theta^\omega(\theta, \xi_i)}_{A_1(\theta, \xi_i)} \times \underbrace{(1 + \text{Frisch el.})}_{B_1(\theta, \xi_i)} \times C_1(\theta, \xi_i)$$

where

$$C_1(\theta, \xi_i) \equiv \int_\theta^{\overline{\theta}} \frac{m_i(\theta)}{m_i(x)} \cdot \frac{u'(c_1(\theta, \xi_i))}{u'(c_1(x, \xi_i))} \cdot \Lambda(\theta, x, \xi_i) \cdot \left[1 - \frac{p_i m_i(x) u'(c_1(x, \xi_i))}{\mu(\xi_i)}\right] \frac{f(x)}{1 - F(\theta)} dx$$

$$\Lambda(\theta, x, \xi_i) \equiv \exp\left(\sigma \int_\theta^x \frac{\eta_i(z)}{z} dz\right)$$

- If $\eta_i(z) > 0 \Rightarrow \Lambda < 1 \Rightarrow$ *reduce* the labor wedge (reward truthful agent).
- If $\eta_i(z) < 0 \Rightarrow \Lambda > 1 \Rightarrow$ *increase* the labor wedge (punish deviating agent).

Back

Calibrating Skill Distribution

- Construct "current policy" economy: agents solve

$$\max u(c_0) - h(\ell_0) + \frac{\beta}{\sigma} \log\left[\sum_i p_i \exp\left(\sigma\left[u\left(c_1(\xi_i)\right) - h\left(\ell_1(\xi_i)\right)\right]\right)\right]$$

  subject to the budget constraints:

$$c_0 + s_1 = \theta\ell_0 - T(\theta l_0)$$

$$c_1(\xi_i) = \omega(M, \theta, \xi_i)\ell_1 - T\left(\omega(M, \theta, \xi_i)\,\ell_1\right) + Rs_1, \forall \xi_i$$

  - CRRA + constant-Frisch w/ typical parameters ⏩ Details
  - HSV-form income tax: $T(y) = y - \zeta y^{1-\tau}$
  - Use Heathcote and Tsujiyama (2021) for (τ, ζ)
- Approx. $F(\theta)$ with mix of normal, lognormal, Pareto
- Target earnings from Guvenen, Pistaferri, and Violante (2022) ⏩ Calibrated distribution of skills
- Set $M = 1778$: corresponds to $F(\theta) = 1 - 10^{-4}$
- Cost of AI ϕ chosen so that *optimal* M w/o ambiguity is 1778. $\phi M/\text{output} = 14\%$.

Parameterizing the Model

- Functional forms:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}$$

$$h(\ell) = \frac{l^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}}$$

- External parameters:
  - $\sigma = 2$
  - $\gamma = 0.5$
  - $\beta = 0.8$ ($R = 1.25$): 15 years at 1.5% real annual return

Back

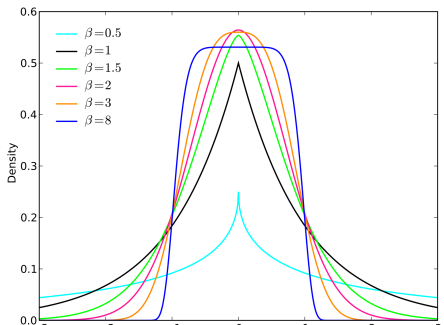
Calibrated Distribution of Skill

Param.	Meaning	Value	Moment	Model	Data
μ_N	Mean of normal	5.24	Mean	6.00	5.62
σ_N	St. dev. of normal	3.65	St. deviation	12.33	12.45
μ_L	Mean of lognormal	0.79	Mode	4.39	4.39
σ_L	St. dev. of lognormal	0.88	10th percentile	0.99	0.96
α	Pareto shape	1.57	75/25 perc. ratio	3.18	3.19
θ^T	Threshold of Pareto	18.17	90/10 perc. ratio	11.34	11.06
ρ	Fraction of Pareto	0.47	Skewness	1.19	1.22
p_N	Probability of normal	0.51	Kurtosis	3.19	3.18
p_L	Probability of lognormal	0.21	Top 1% share	11.68%	11.44%

Back

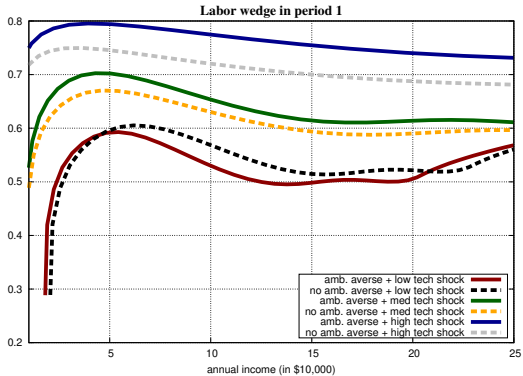
More on Generalized Normal Distribution

- We assume *generalized normal distribution* for $\ln(\xi - 1)$
  - Shape parameter: $\beta = 1, 2, 3$; lower $\beta \Rightarrow$ fatter tails



Back

Labor Distortion in t=1



(a) Labor wedges in $t = 1$

- Deviating agent places *more* weight on high-tech shocks than the truthful agent
  - ⇒ *Increase* labor wedge in "good" times.
- Deviating agent place *less* weight on low-tech shocks than the truthful agent.
  - ⇒ *Decrease* labor wedge in "bad" times.

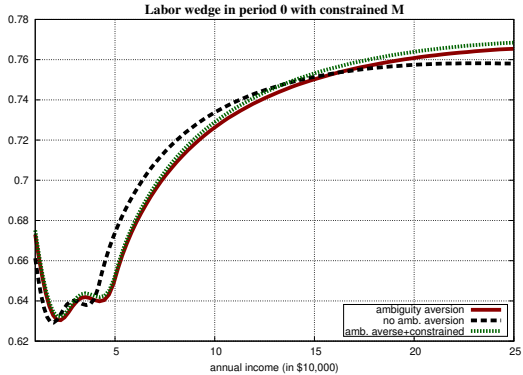
Back

Optimal AI Investment

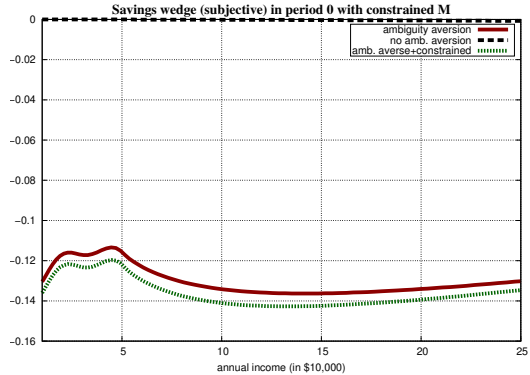
	Shape parameter for ξ:		
	2	**4** (baseline)	**6**
$\eta = 0.15$ (ambiguity aversion)	53.1	67.5	70.6
$\eta \to \infty$ (no ambiguity aversion)	84.5	100.0	102.5

- Optimal AI investment is lowered by 32% $\Rightarrow \phi M/\text{output} = 8.4\%$.
- AI investment falls more as extreme shocks become more likely
- ⏩ What if we keep AI fixed

Keeping M fixed



(a) Period 0 labor wedge



(b) Period 0 savings wedge

Labor wedge goes up slightly, savings subsidy increases. Back