Inequality Sensitive Optimal Treatment Assignment

Eduardo Zambrano, Cal Poly

This material is based upon work supported by the NSF under Grant # 2313969

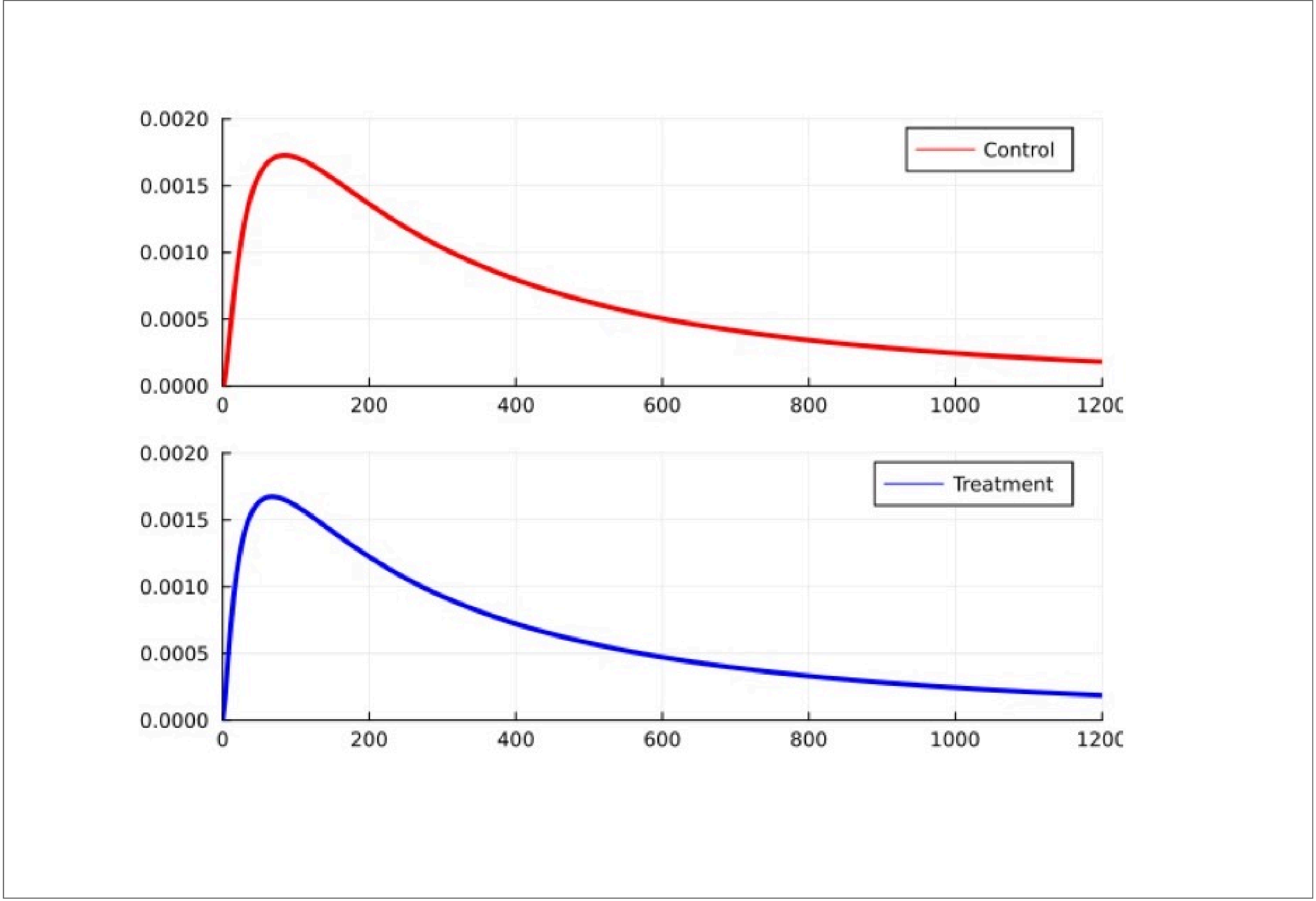
Outline of a research program

- Project 1: **Protected Income and Inequality Aversion**
- Project 2: Social Preferences Under Uncertainty and Ambiguity
- Project 3: `Inequality Sensitive Optimal Treatment Assignment` (This paper)
- Egalitarian Equivalent Treatment Effect Estimation
- Inequality Aversion Elicitation in Theory and in Practice

The main preliminary issues

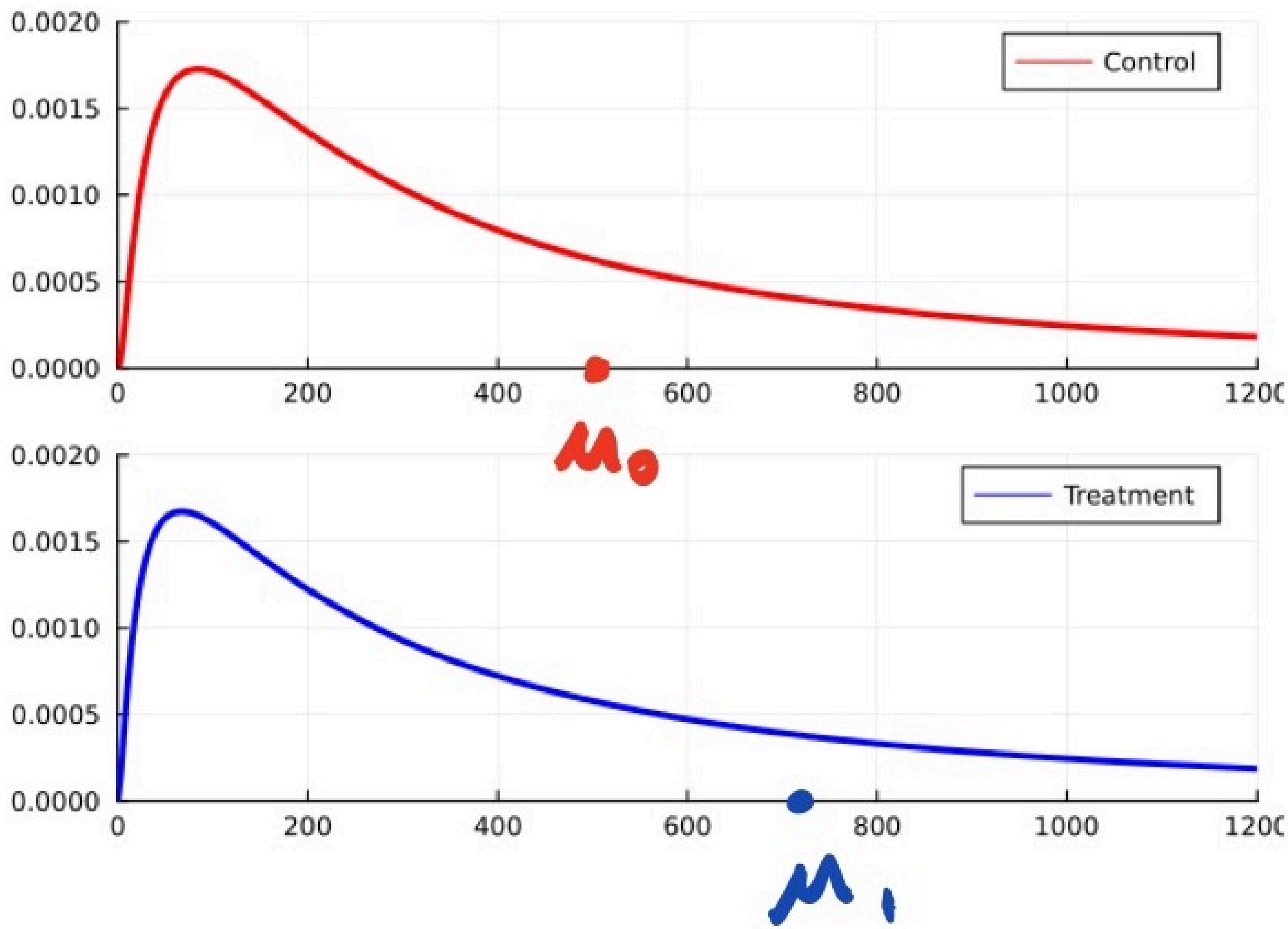
- How to summarize HTE
- How to account for magnitude of losses when errors occur
- How to properly embed the answers to the questions above in an econometric framework

Two treatments and their outcome distributions



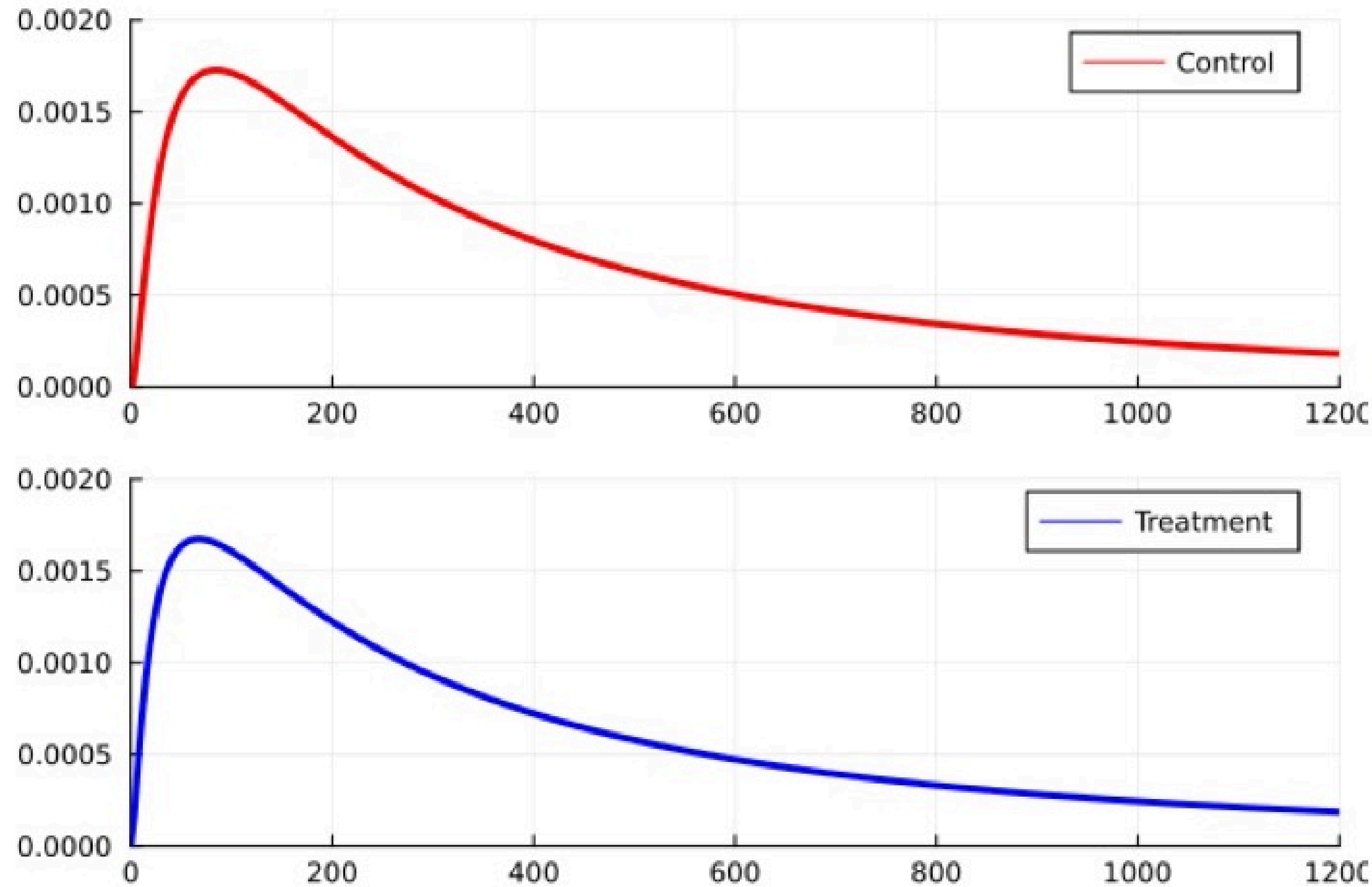
0.0020
0.0015
0.0010
0.0005
0.0000
0
200
400
600
800
1000
120C
Control
Treatment

An inequality neutral evaluator



0.0020
0.0015
0.0010
0.0005
0.0000
0
200
400
600
800
1000
1200
Control
Treatment

Incorporating inequality aversion

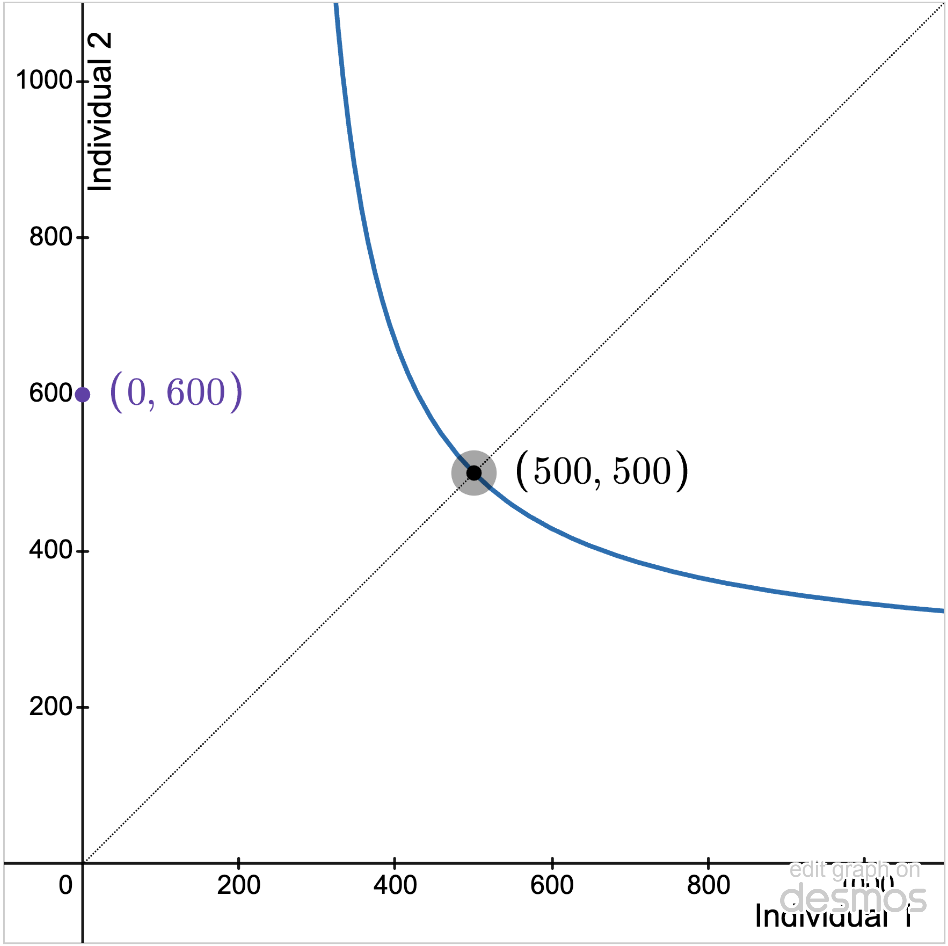


0.0020
0.0015
0.0010
0.0005
0.0000
0
200
400
600
800
1000
1200
Control
Treatment

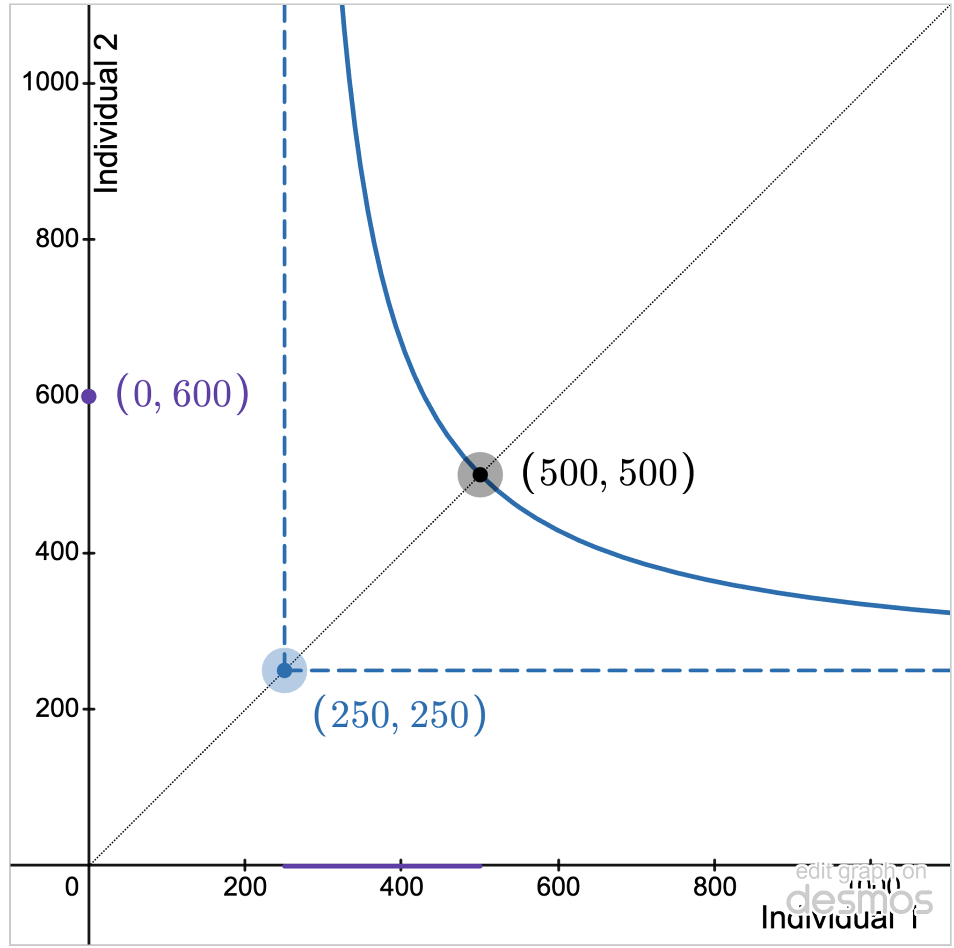
Social Preferences

- $W\left(y_1, \ldots, y_n\right) = \sum_{i=1}^{n} f\left(y_i\right)$

Investigate acceptable tradeoffs



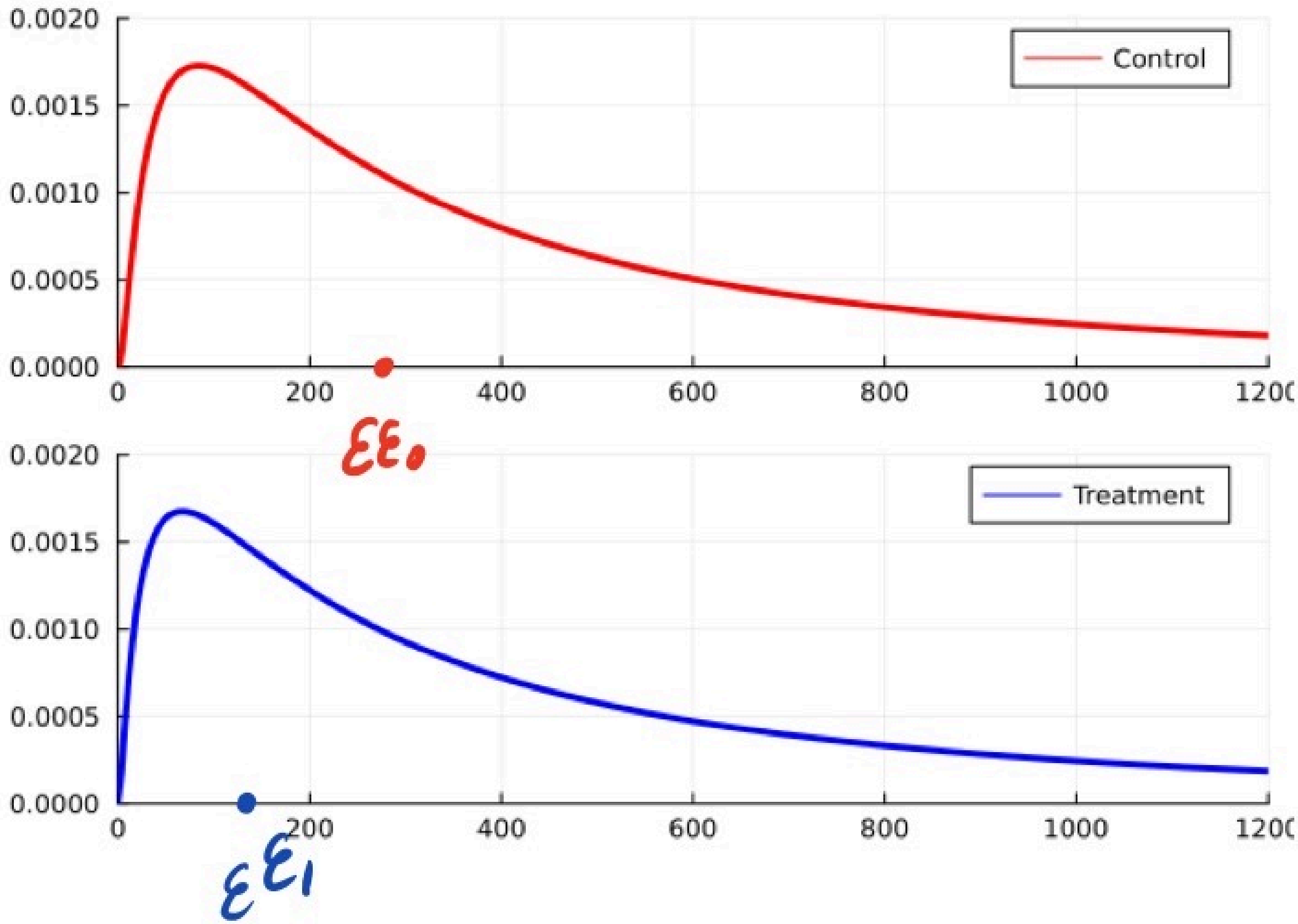
Notice the asymptotes



From Project 1

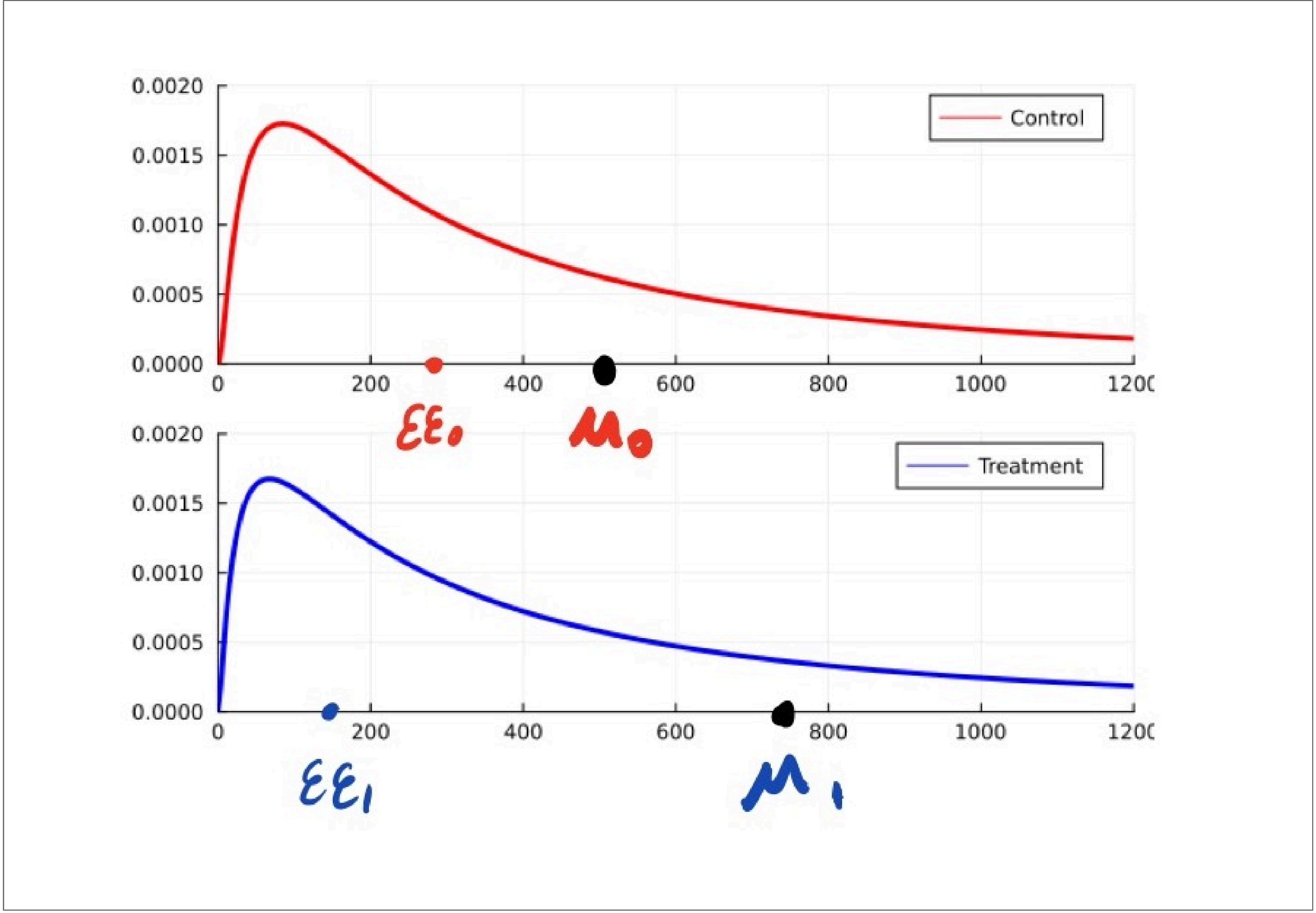
Theorem 1 (With Marc Fleurbaey) If you know where the asymptotes lie, then you know f.

Egalitarian Equivalent ($\mathcal{EE}$)



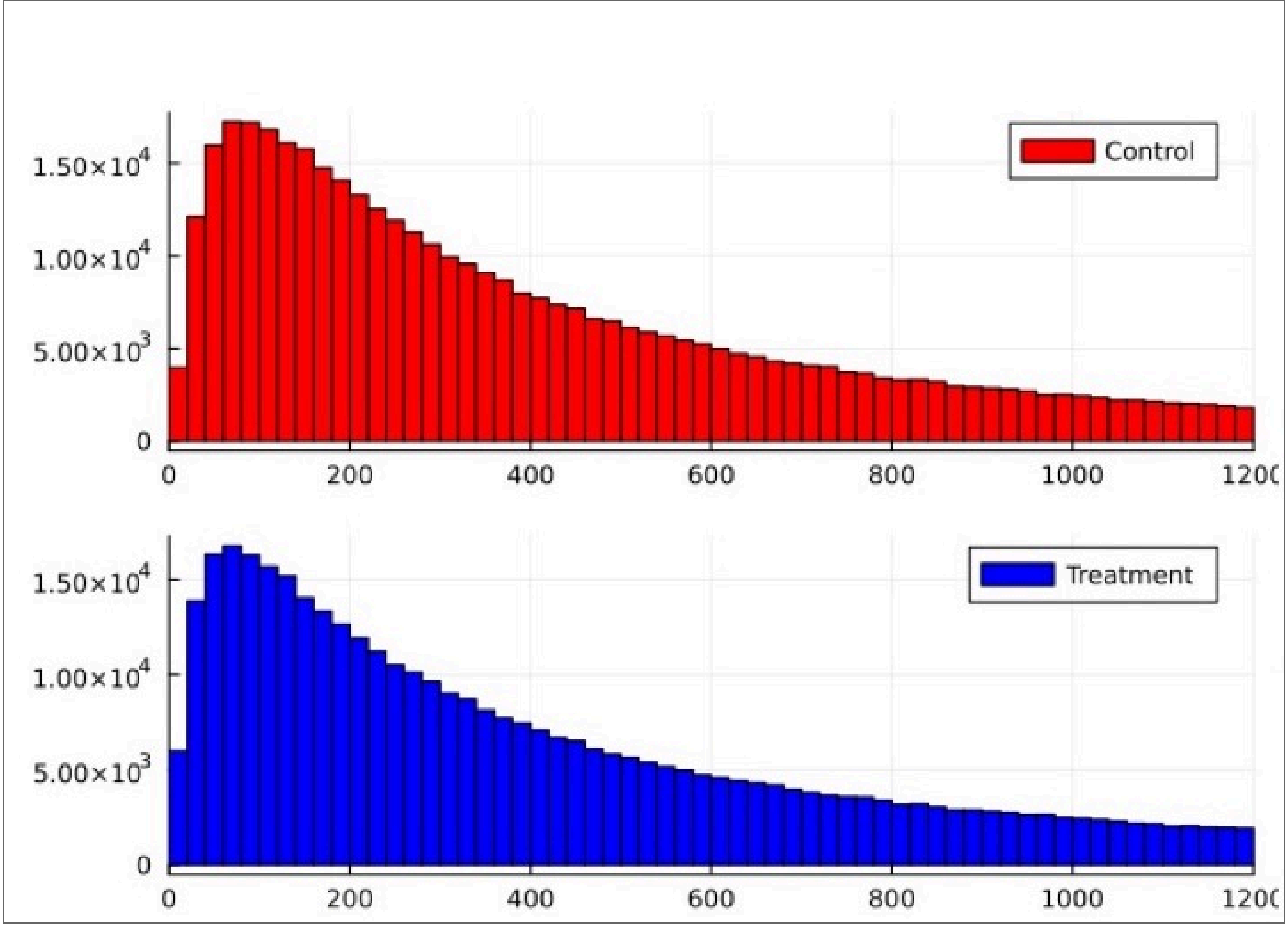
0.0020
0.0015
0.0010
0.0005
0.0000
0
200
400
600
800
1000
1200
Control
$\varepsilon\varepsilon_0$
Treatment
$\varepsilon^{\varepsilon_1}$

Inequality Aversion

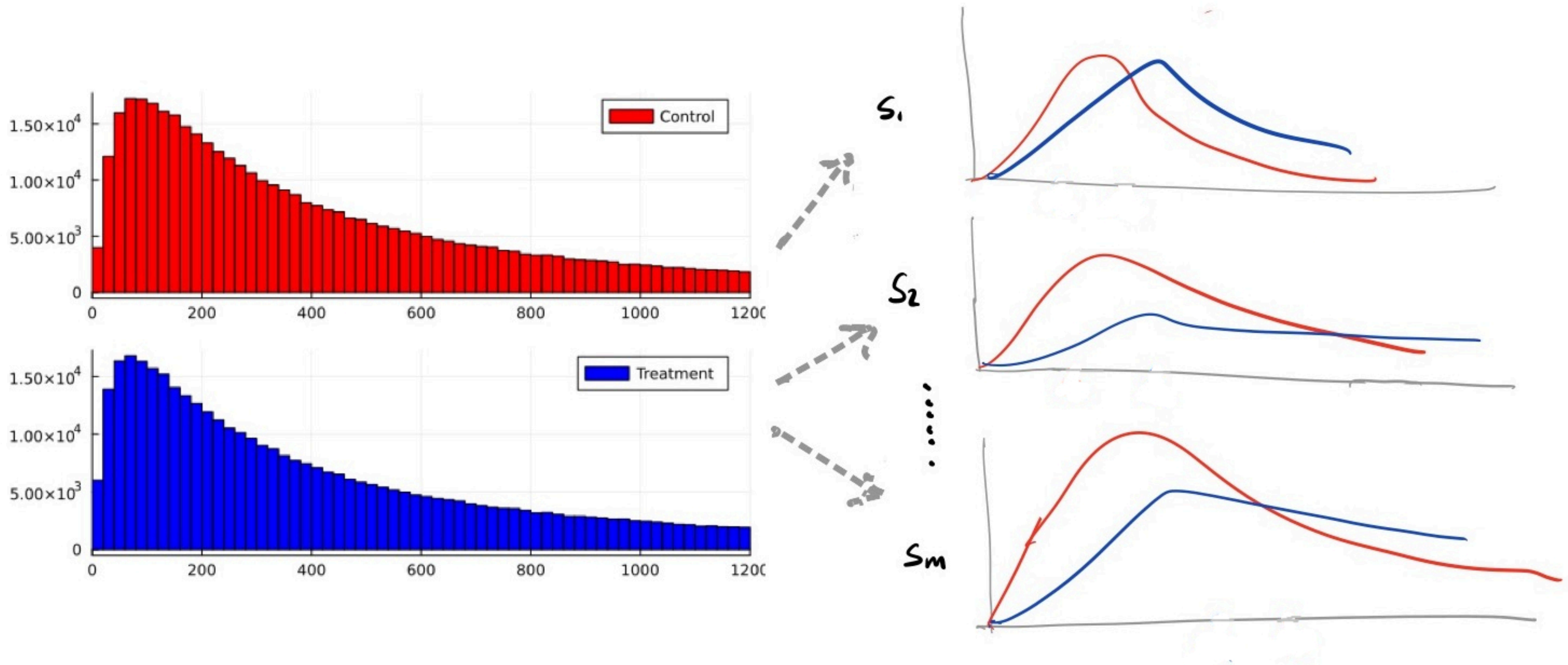


Control
Treatment
0.0020
0.0015
0.0010
0.0005
0.0000
0
200
400
600
800
1000
$\varepsilon\varepsilon_0$
μ_0
$\varepsilon\varepsilon_1$
μ_1

Incorporating uncertainty


Control
Treatment
1.50×10^4
1.00×10^4
5.00×10^3
0
0
200
400
600
800
1000
1200

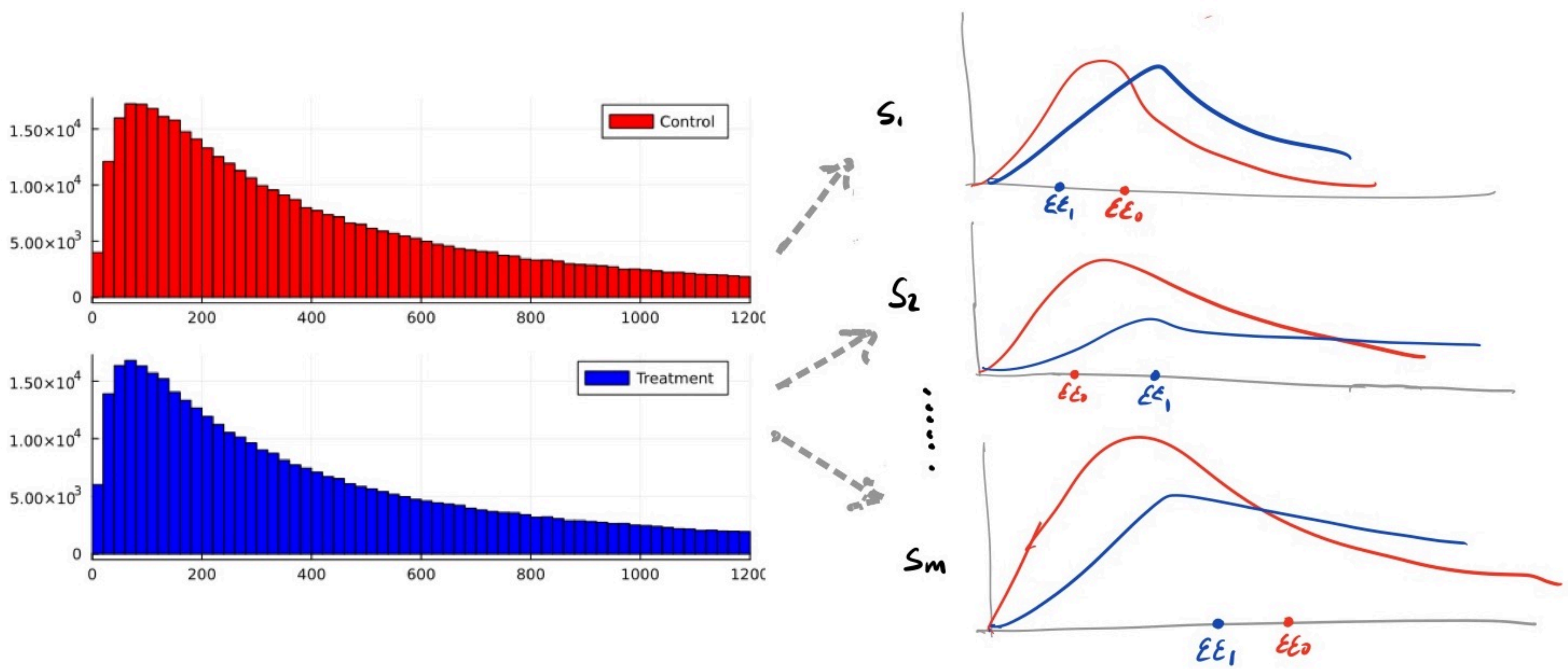
Incorporating uncertainty



From Project 2

Theorem 2 To avoid conflating risk aversion and inequality aversion, use $\mathcal{EE}$.

Identify each treatment with the profile $(\mathcal{E}\mathcal{E}^1, \ldots, \mathcal{E}\mathcal{E}^m)$



An illustration of Theorem 2

Click here

This Paper (Project 3)

- Egalitarian Equivalent Optimal Statistical Decisions
  - Under point and partial identification
  - Small and large sample analysis
  - Applications: Microcredit, JobCorps

Statistical Decision Theory

- $d \in \{a, b\}$
- $y^s(d)) = \begin{bmatrix} y_1^s(d) \\ \vdots \\ y_i^s(d) \end{bmatrix}$
- $ee^s(d) := ee(y^s(d)) = f^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} f(y_i^s(d))\right)$
- f is determined using **Theorem 1**

Egalitarian Equivalent Treatment Effect (EETE)

$$\tau(s) := ee^s(b) - ee^s(a)$$

Statistical Treatment Rules

- $\delta : \mathcal{P} \to [0, 1]$

- $\delta : \mathcal{P} \to [0,1]$

$$ee^s(\delta(P_s)) =$$

$$f^{-1}\left[\delta(P_s) f(ee^s(b)) + (1-\delta(P_s)) f(ee^s(a))\right]$$

$$ee^s(\delta(P_s)) = f^{-1}\left[\delta(P_s)\left(\frac{1}{n}\sum_{i=1}^{n} f\left(y_{is}(b)\right)\right) + (1-\delta(P_s))\left(\frac{1}{n}\sum_{i=1}^{n} f\left(y_{is}(a)\right)\right)\right]$$

Thanks to **Theorem 2**, the decision problem boils down to selecting δ in order to obtain the most favorable profile

$$(ee^1(\delta(P_s)), \ldots, ee^m(\delta(P_s)))$$

according to the preferences $\succeq$ over the restricted set Υ_1, and given knowledge of the sampling distribution P_s.

- Let π be a prior on S
- for each P_s, let $S(P_s) \subseteq S$ denotes the state space obtained with knowledge of P_s.

The Bayes Decision Problem

$$\max_{\delta(P_s)\in[0,1]} E_\pi\left[ee^s(\delta(P_s))|S(P_s)\right]$$

$$(ee^1(\delta(P_s)), \ldots, ee^m(\delta(P_s)))$$

The Maximin Decision Problem

$$\max_{\delta(P_s)\in[0,1]} \min_{s\in S(P_s)} ee^s(\delta(P_s))$$

$$(ee^1(\delta(P_s)),\ldots,ee^m(\delta(P_s)))$$

The Minimax Regret Decision Problem

$$\min_{\delta(P_s)\in[0,1]} \max_{s\in S(P_s)} \left[\max\{ee^s(a), ee^s(b)\} - ee^s(\delta(P_s))\right]$$

$$(ee^1(\delta(P_s)), \ldots, ee^m(\delta(P_s)))$$

Bayesian

$$\max_{\delta(P_s)\in[0,1]} E_\pi\left[\boxed{ee^s(\delta(P_s))} \middle| S(P_s)\right]$$

Maximin

$$\max_{\delta(P_s)\in[0,1]} \min_{s\in S(P_s)} \boxed{ee^s(\delta(P_s))}$$

Minimax Regret

$$\min_{\delta(P_s)\in[0,1]} \max_{s\in S(P_s)} \left[\max\{ee^s(a), ee^s(b)\} - \boxed{ee^s(\delta(P_s))}\right]$$

Let s_w, s_a and s_b be such that
$ee^{s_w}(a) = ee^{s_b}(a) = \min_{s \in S(P_s)} ee^s(a)$
$ee^{s_w}(b) = ee^{s_a}(b) = \min_{s \in S(P_s)} ee^s(b)$
and
$ee^{s_d}(d) = \max_{s \in S(P_s)} ee^s(d)$ for $d \in \{a, b\}$.

Theorem 3 Assume that the true state s is partially identified, the set $\{(ee^s(a), ee^s(b))\}_{s\in S(P_s)}$ is bounded, and that $s_w, s_a, s_b \in S(P_s)$. Then

- The solution to the Bayesian decision problem is

$$\delta^B(P_s) = 1\left(E_\pi\left[\tau_{ee}(s)|S(P_s)\right] > 0\right)$$

Theorem 3 (cont.)

- The solution to the maximin decision problem is

$$\delta^M(P_s) = 1\left(\tau_{ee}(\boldsymbol{s}_w) > 0\right)$$

THEOREM 3 (CONT.)

- The solution to the minimax regret decision problem is $\delta^R(P_s) \in (0,1)$ such that

$$ee^{s_a}(a) - ee^{s_a}(\delta^R(P_s)) = ee^{s_b}(b) - ee^{s_b}(\delta^R(P_s))$$

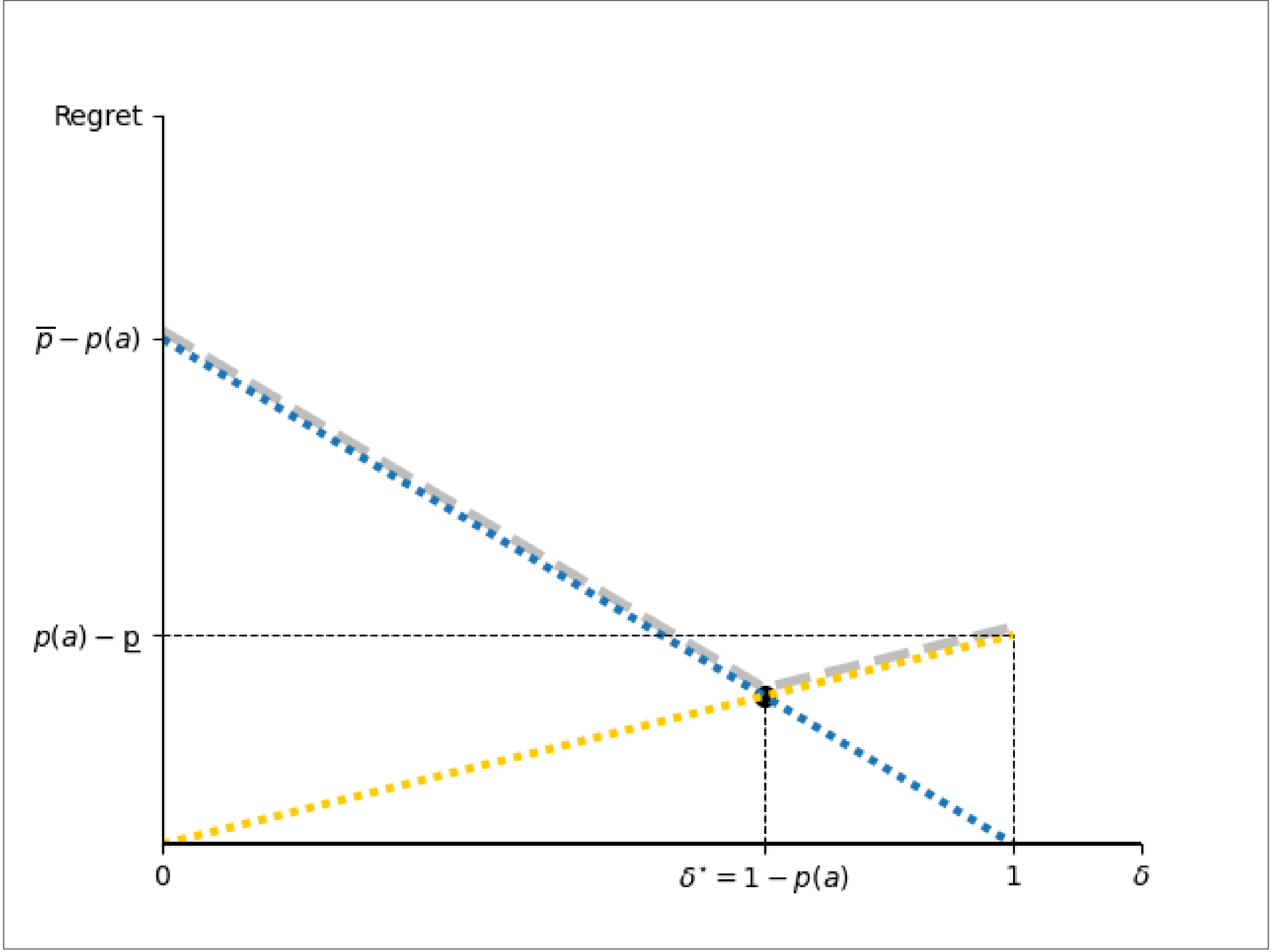
Side by Side

$$\delta^{B}(P_s) = 1\left(E_\pi\left[ee(b)|S(P_s)\right] - E_\pi\left[ee(a)|S(P_s)\right] > 0\right)$$

$$\delta^{M}(P_s) = 1\left(ee^{s_w}(b) - ee^{s_w}(a) > 0\right)$$

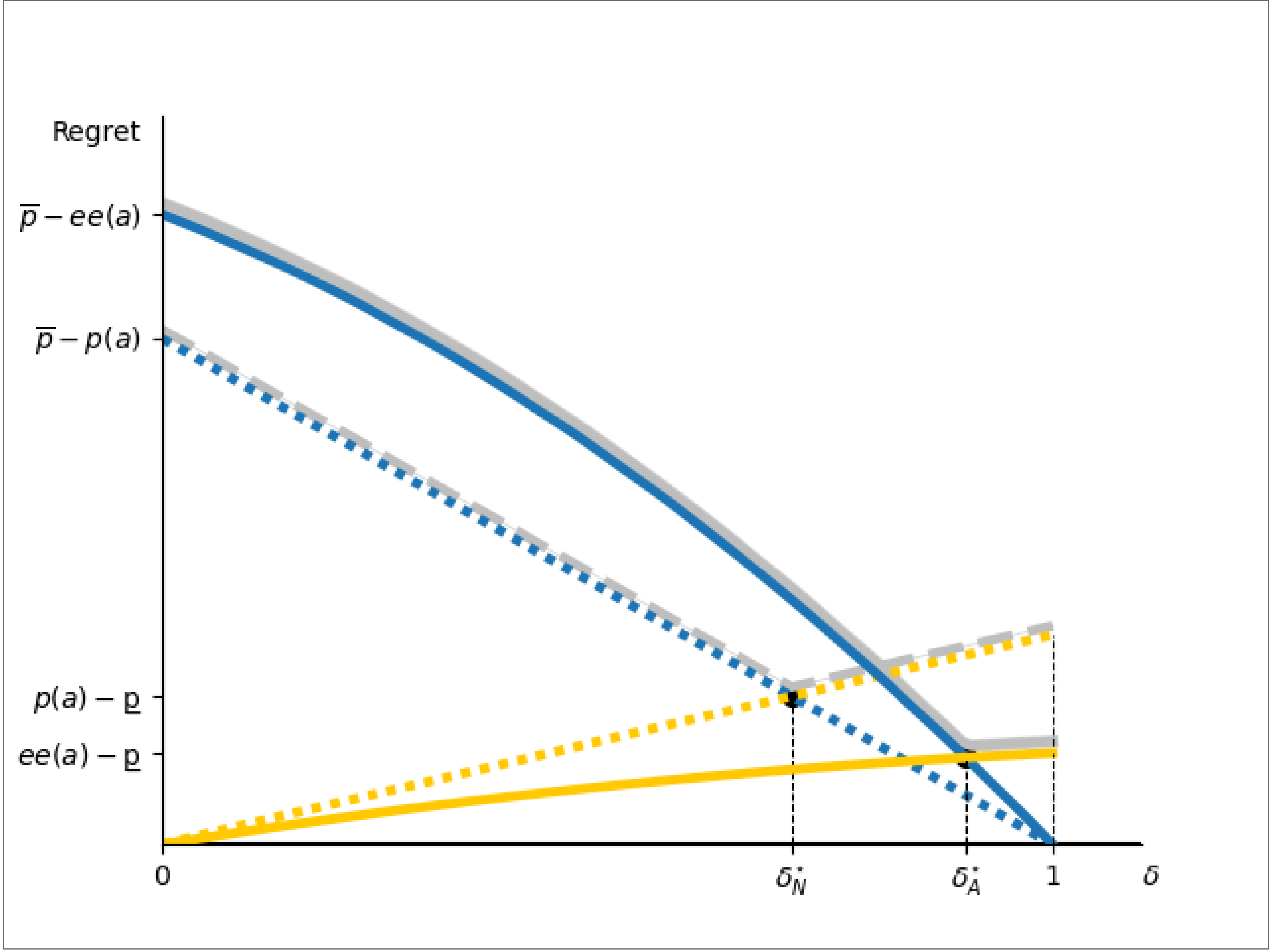
$$ee^{s_a}(a) - ee^{s_a}(\delta^{R}(P_s)) = ee^{s_b}(b) - ee^{s_b}(\delta^{R}(P_s))$$

Minimizing Inequality Neutral Worst Regret



Regret
$\overline{p} - p(a)$
$p(a) - \underline{p}$
0
$\delta^* = 1 - p(a)$
1
δ

Minimizing Inequality Averse Worst Regret



Regret
$\overline{p} - ee(a)$
$\overline{p} - p(a)$
$p(a) - \underline{p}$
$ee(a) - \underline{p}$
0
δ_N^*
δ_A^*
1
δ

Application: A Bayesian Meta Analysis of the Microcredit Literature

- Meager 2022 estimates posterior distributions of the effect of microcredit interventions on consumption using data from randomized trials that expand access to microcredit in five countries.

Meager’s results

- Meager reports considerable treatment effect heterogeneity.
- Large segments of the distribution of consumption are nearly unaffected by the policy (from the 5-th to the 75-th percentiles)
- This is coupled with with large yet uncertain differences on the upper tails of the distribution of consumption of the treatment and control groups, especially within the group of households with previous business experience.
- Given that the treatment will probably increase inequality, “the social welfare effects of microcredit are likely to be complex.” (Meager 2022, p. 1821).

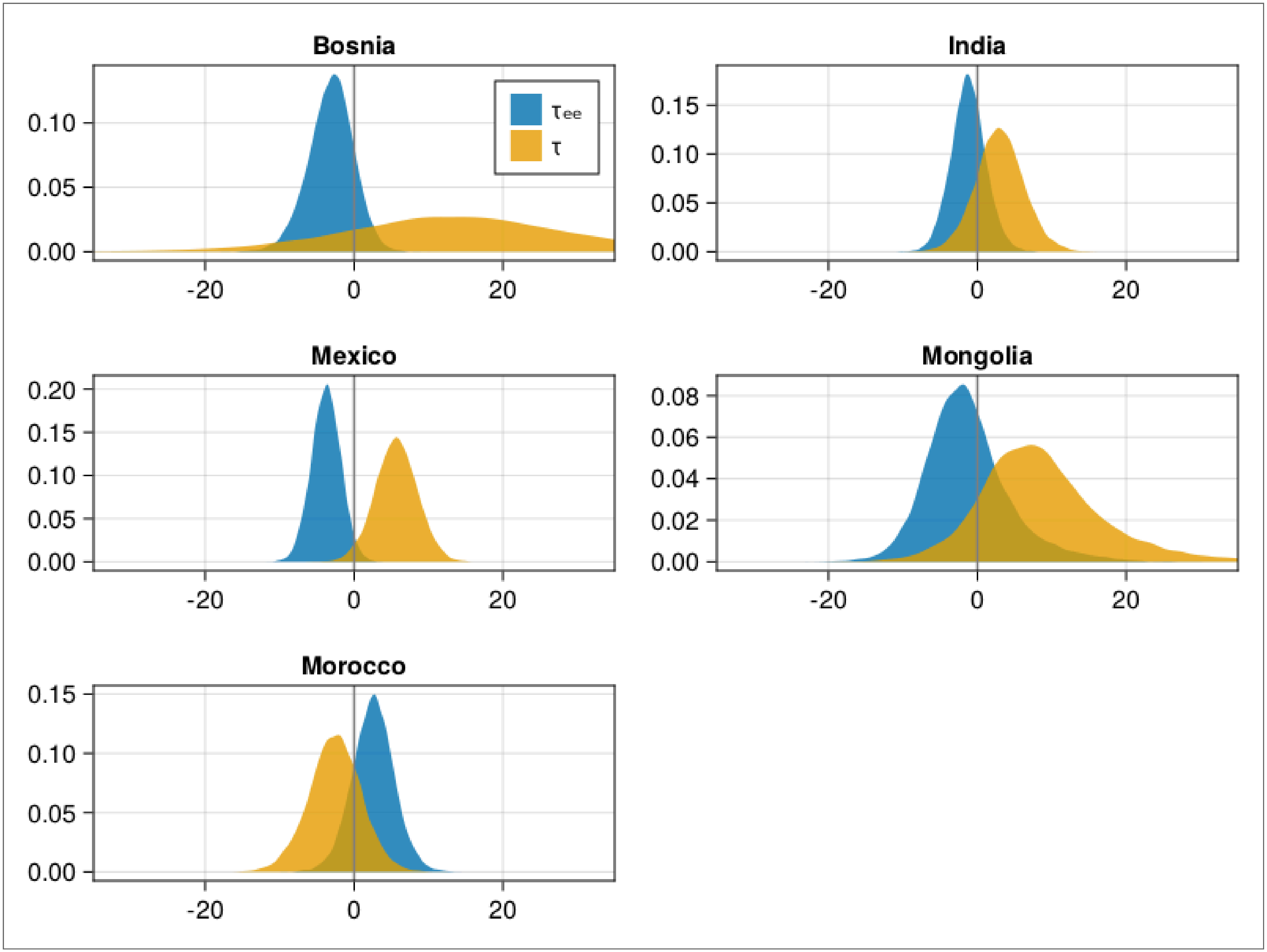
Meager's Bayesian hierarchical model

$$y_{ik}(T_{ik}) \sim \texttt{LogNormal}(\mu_k + \zeta_k T_{ik}, \sigma_k \lambda_k^{T_{ik}}) \text{ for } k = 1, \ldots, 5;$$

$$0.1\mu_k, 0.1\zeta_k, \log(\sigma_k), \log(\lambda_k) \sim \texttt{MvN}(0, 10I) \text{ for } k = 1, \ldots, 5.$$

- I compute mean treatment effects and egalitarian equivalent treatment effects using Meager's Markov Chain Monte Carlo (MCMC) output, denoted $\hat{\pi}$, which contains three chains with four thousand draws per chain.
- The egalitarian equivalent at every draw can be computed using the expression

$$\mathcal{E}\mathcal{E}(y_k(d)) = e^{\mu_k + \zeta_k d + \frac{1}{2}(1-\gamma)\left(\sigma_k \lambda_k^d\right)^2}.$$



Bosnia
India
Mexico
Mongolia
Morocco
τ_{ee}
τ
0.00
0.05
0.10
0.15
0.20
0.02
0.04
0.06
0.08
-20
0
20

Posterior distributions of the mean treatment effects (τ) and the egalitarian equivalent treatment effects (τ_{ee}). All units are 2009 USD PPP per two weeks.

Microcredit treatment effect Bayesian estimates

	$E_{\hat{\pi}}[\tau]$	$P_{\hat{\pi}}[\tau > 0]$	$E_{\hat{\pi}}[\tau_{ee}]$	$P_{\hat{\pi}}[\tau_{ee} > 0]$
Bosnia	13.82	82.5%	-3.11	14.3%
India	3.00	83.1%	-1.20	29.1%
Mexico	5.65	97.5%	-3.76	2.7%
Mongolia	8.58	88.1%	-1.35	35.0%
Morocco	-2.57	22.7%	2.55	82.4%

Quantile:	5th	25th	35th	45th	55th	65th	75th	95th
Partial pooling								
Bosnia	−5.2	−3.8	−3.9	−3.6	−2.8	−1.1	2.6	52.4
	(−9.6,2.6)	(−9,0.7)	(−9.9,1)	(−10.9,1.7)	(−12.4,3.6)	(−14.7,8.6)	(−19.4,20.9)	(−75.8,188.3)
India	−2	−1.2	−0.6	0.1	1.1	2.4	4.3	16
	(−5.3,1.4)	(−4.9,2.7)	(−4.7,3.6)	(−4.4,4.9)	(−4.2,6.6)	(−4,9)	(−4,12.8)	(−5.6,37.9)
Mexico	−4.7	−3.4	−2.2	−0.8	1.2	3.9	8	34.1
	(−7.3,−2.1)	(−6.5,−0.3)	(−5.8,1.2)	(−4.7,3)	(−3.5,5.6)	(−1.7,9.3)	(0.7,15)	(15.5,52.7)
Mongolia	−3	−1.7	−0.6	0.7	2.7	5.8	10.3	38.4
	(−11.4,5.3)	(−9.2,9.8)	(−7.5,12.2)	(−6,15.8)	(−5.3,20.2)	(−5.7,26.5)	(−7.3,36.2)	(−22.4,108)
Morocco	4.3	2.9	2	0.9	−0.4	−2.2	−4.6	−18.8
	(−0.5,9)	(−2.1,7.8)	(−3.1,7.1)	(−4.6,6.4)	(−6.7,5.6)	(−9.6,5)	(−14,4.4)	(−41.5,3.3)

Notes: All units are US$ PPP per two weeks. Estimates are shown with their 95 percent uncertainty intervals below them in parentheses.

Table 1: Bayesian quantile treatment effects on consumption (from Table 1 in Meager 2022)

EETE estimation workflow

1. Use **Theorem 1** to help determine which social preference under certainty to bring into the analysis.
2. Use **Theorem 2** to decide how to incorporate that social preferences into a world where risk, uncertainty or ambiguity play a prominent role.
3. Use **Theorem 3** to identify the correct optimal statistical treatment rule needed for the problem at hand.

Thank You For Coming!

Appendix 1: Protected Income and Inequality Aversion

- Consider a population of n individuals
- Each individual i has a known income $y_i \geq 0$.

Social Preferences

- $W\left(y_1, \ldots, y_n\right) = \sum_{i=1}^{n} f\left(y_i\right)$
- Let $y_1(y_2, y)$ solve

$$f(y_1(y_2, y)) + f(y_2) = 2f(y)$$

Protected Income

$$\ddot{Y}\left(y\right):=\inf\left\{y_1\left(y_2,y\right)\right|\left(y_2,y\right)\in D_2\}.$$

Kolm - Atkinson (K-A)

$$f(y) = \begin{cases} \frac{y^{1-\gamma}}{1-\gamma} & \gamma \neq 1, \gamma \geq 0 \\ \ln y & \gamma = 1 \end{cases}$$

How much Inequality Aversion is not enough?

- $\ddot{Y}(y) = 2^{\frac{1}{1-\gamma}} y$ for $\gamma > 1$
- $\ddot{Y}(y) = 0$ for $\gamma \leq 1$

Protected Income (against n)

- $y_1(y_2, y_3, \ldots, y_n, y)$ is defined implicitly by the equation $f(y_1(y_2, y_3, \ldots, y_n, y)) + f(y_2) + \cdots + f(y_n) = nf(y)$.

  In the case of the *K-A* class, one obtains

$$\mathring{Y}(y) := \inf \{y_1(y_2, y_3, \ldots, y_n, y) | (y_2, y_3, \ldots, y_n, y) \in D_n\}$$

$$= n^{\frac{1}{1-\gamma}} y$$

Theorem 1 (With Marc Fleurbaey) The Kolm-Atkinson class of SPs with $\gamma > 1$ is (up to an affine transform) characterized by the requirement that there exist $\lambda, \mu \in (0,1)$ such that for all $y > 0, \ddot{Y}(y) = \lambda y$ and $\mathring{Y}(y) = \mu y$.

One must then have $\lambda = 2^{\frac{1}{1-\gamma}}$ and $\mu = n^{\frac{1}{1-\gamma}}$ for some $\gamma > 1$.

Appendix 2: Welfare Economics Under Uncertainty and Ambiguity

An extension of Theorem 1 in Fleurbaey (2010).

- $N = \{1, \ldots, n\}$
- $S = \{s_1, \ldots, s_m\}$
- $y = (y_i^s)_{i \in N, s \in S}$
- $\Upsilon \subseteq \mathbb{R}^{nm}$
- P, R, I: the evaluator's preferences over Υ
- $[y^s] \in \Upsilon^c$
- $(y_i) \in \Upsilon^e$
- $\succ, \succeq, \sim$: the evaluator's preferences over Υ^1

Continuity

Let $y, y^{'} \in \Upsilon$ and $(y(t))_{t\in\mathbb{N}} \in \Upsilon^{\mathbb{N}}$ be such that $y(t) \to y$. If $y(t)Ry^{'}$ for all $t \in \mathbb{N}$, then $yRy^{'}$. If $y^{'}Ry(t)$ for all $t \in \mathbb{N}$, then $y^{'}Ry$.

Weak Dominance

For all $y, y^{'} \in \Upsilon$, one has $yRy^{'}$ if for all $s \in S$, $[y^s]R[y^{s'}]$.

Weak Pareto for No Risk

For all $[y^s], [y^{'s}] \in \Upsilon^c$, one has $[y^s]P[y^{'s}]$ if, for all $i \in N$, $y_i^s > y_i^{'s}$.

Statistical Extension

For all $(y_i), (y_i^{'}) \in \Upsilon^e$, one has $(y_i)P(y_i^{'})$ if $y_i \succ y_i^{'}$.

Let $\mathcal{E}\mathcal{E}(y^s)$ be the continuous function such that, for each $y^s \in \mathbb{R}^n$,

$$[y^s] \; I \; [(\mathcal{E}\mathcal{E}(y^s), \ldots, \mathcal{E}\mathcal{E}(y^s))].$$

Theorem 2 Let $\Upsilon = \mathbb{R}^{nm}$ and R satisfy Continuity, Weak Dominance, Weak Pareto for No Risk, and Statistical Extension. For all $y, y^{'} \in \Upsilon$, one has

$$yPy^{'}$$

$$iff$$

$$(\mathcal{E}\mathcal{E}(y^1), \ldots, \mathcal{E}\mathcal{E}(y^m)) \succ (\mathcal{E}\mathcal{E}(y^{'1}), \ldots, \mathcal{E}\mathcal{E}(y^{'m}))$$

Table 2: Prospects

	State 1	State 2
y_1	4	2
y_2	5	2
y_3	6	2

	State 1	State 2
y_1	3	1
y_2	8	1
y_3	10	1

$\pi = (1/3, 2/3)$

Prospect 1	State 1	State 2
y_1	4	2
y_2	5	2
y_3	6	2
Geometric mean (across indiv.)	4.93	2
E[Geometric mean] (across states)	3.95	

Prospect 2	State 1	State 2
y_1	3	1
y_2	8	1
y_3	10	1
Geometric mean (across indiv.)	6.21	1
E[Geometric mean] (across states)	4.48	

Prospect 1	State 1	State 2
$ln(y_1)$	1.31	0.69
$ln(y_2)$	1.61	0.69
$ln(y_3)$	1.79	0.69
Arithmetic mean of logs (across indiv.)	1.60	0.69
E[Arithmetic mean of logs] (across states)	1.29	

Prospect 2	State 1	State 2
$ln(y_1)$	1.1	0
$ln(y_2)$	2.08	0
$ln(y_3)$	2.30	0
Arithmetic mean of logs (across indiv.)	1.83	0
E[Arithmetic mean of logs] (across states)	1.22	

Go Back

Appendix 3: A Differences in Means comparison

- The point estimates (with standard errors in parentheses) of the differences in means $E[y(1)] - E[y(0)]$ for the five sites are: Bosnia -1.59 (14.14), India 4.55 (3.85) (India), Mexico 5.51 (2.90), Mongolia 50.45 (15.67) and Morocco -2.93 (4.26).
- One would reject the hypothesis that the treatments have the same effect on average income at the 5% level in the case of Mongolia, and would not reject the hypothesis in the other four countries.

Appendix 4: Discussion

- Interventions that exhibit considerable treatment effect heterogeneity can be difficult to evaluate and summarize in terms of welfare.
- The methods I propose aim to provide specific, quantitative guidance for how to do this evaluation.
- The microcredit empirical illustration I highlight above shows that performing the evaluation in an inequality sensitive way can make a difference in terms of how one ranks treatments, relative to when inequality considerations are put aside.

Appendix 5: Minimizing Egalitarian Equivalent Regret at JobCorps

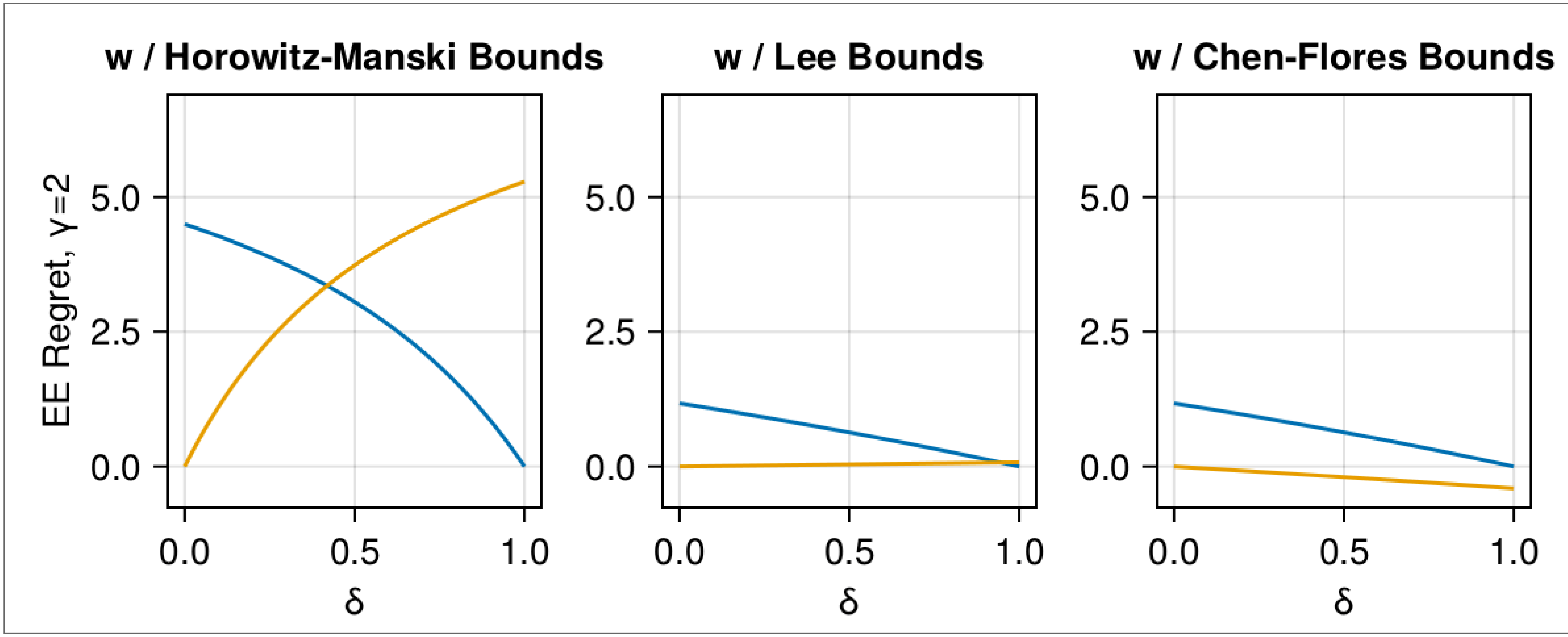
- JobCorps is a widely studied RCT education and training program for disadvantaged youth.
- The evaluator only observes wages for those individuals who are employed.
- $ee^s(a)$ and $ee^s(b)$ are not point identified and therefore $\tau_{ee}(s)$ is also not point identified at s.
- Partial identification is achievable under relatively mild assumptions.

Egalitarian Equivalent Bounds

$$\gamma = 2$$

	$ee^L(a)$	$ee^U(a)$	$ee^L(b)$	$ee^U(b)$
Horowitz and Manski	4.1	9	3.8	8.6
Lee	6.6	6.6	6.5	7.7
Chen and Flores	6.6	6.6	6.8	7.7

Minimizing Egalitarian Equivalent Regret at JobCorps



Optimal treatment assignment

$$\gamma = 2$$

	Horowitz and Manski Bounds	Lee Bounds	Chen and Flores Bounds
δ^*	0.42	0.95	1

$y_1(y_2, y)$ solves

$$f(y) = \begin{cases} \frac{y^{1-\gamma}}{1-\gamma} & \gamma \neq 1, \gamma \geq 0 \\ \ln y & \gamma = 1 \end{cases}$$

$$f(y_1(y_2, y)) + f(y_2) = 2f(y)$$