Learning Treatment Policies from Observational Data

Mike Oberst
MIT

moberst.com

Get the slides here!

Models, Inference, and Algorithms (MIA) Primer

Overview

Goals

- By the end of this primer, you should understand…
  - What a **causal effect** is, and when we can infer it from observational data

    "What is the effect of treatment vs. placebo?"
  - How **causal effect estimation** is a special case of **policy evaluation**

    "What would happen if I implemented a new set of guidelines?"
  - How to use policy evaluation to **learn optimal policies**

    "What is the optimal treatment strategy?"

Goals

- By the end of this primer, you should understand…
  - What a **causal effect** is, and when we can infer it from observational data

    "What is the effect of treatment vs. placebo?"
  - How **causal effect estimation** is a special case of **policy evaluation**

    "What would happen if I implemented a new set of guidelines?"
  - How to use policy evaluation to **learn optimal policies**

    "What is the optimal treatment strategy?"

Causal Effects

(and related quantities, like the Average Treatment Effect)

What is a causal effect?

- What is the effect of…
  - **Thiazide** on **cardiac outcomes** in **hypertensive patients**? [1]
  - **Hydroxychloroquine** on **adverse events** in **COVID-19** patients? [2]
  - …
- **Causal questions** from large **observational** datasets

[1] M. A. Suchard et al., "Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes : a systematic , multinational, large-scale analysis," Lancet, vol. 6736, no. 19, pp. 1–11, 2019.
[2] J. C. E. Lane et al., "Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study," medRxiv, 2020.

Conclusions of an Observational Study

THE LANCET

ARTICLES | VOLUME 394, ISSUE 10211, P1816-1826, NOVEMBER 16, 2019

Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis

Prof Marc A Suchard, MD • Martijn J Schuemie, PhD • Prof Harlan M Krumholz, MD • Seng Chan You, MD • RuiJun Chen, MD
Nicole Pratt, PhD • et al. Show all authors

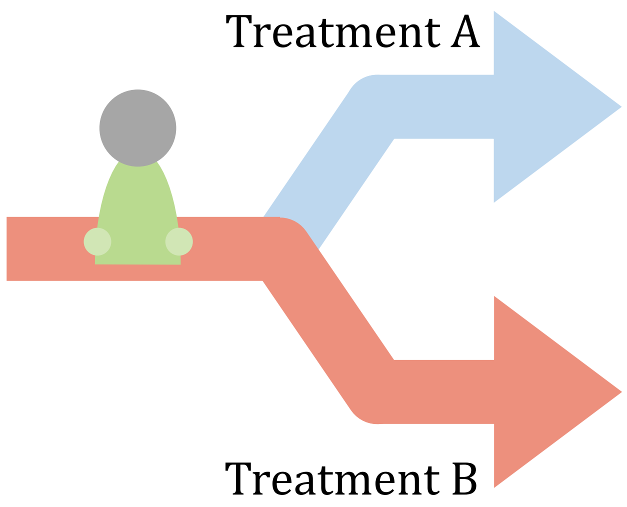
Published: October 24, 2019 • DOI: https://doi.org/10.1016/S0140-6736(19)32317-7 • Check for updates

*"**thiazide** [...] showed **better primary effectiveness** than **[ACE] inhibitors**" [1]*

[1] M. A. Suchard et al., "Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes : a systematic , multinational, large-scale analysis," Lancet, vol. 6736, no. 19, pp. 1–11, 2019.

"What is the effect" as a causal question

Every patient has a set of potential outcomes $Y(t)$, corresponding to different treatment decisions t.



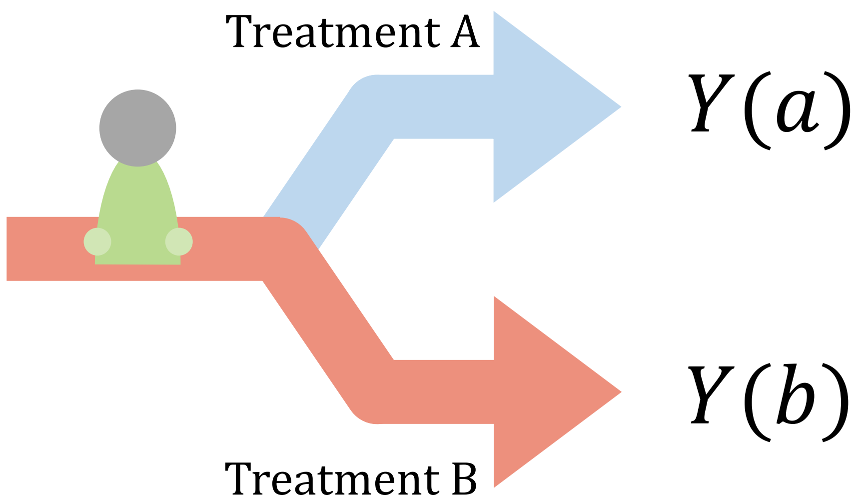
$Y(a)$

$Y(b)$

Average Treatment Effect: $E[Y(a) - Y(b)]$

"What is the effect" as a causal question

Every patient has a set of potential outcomes $Y(t)$, corresponding to different treatment decisions t.



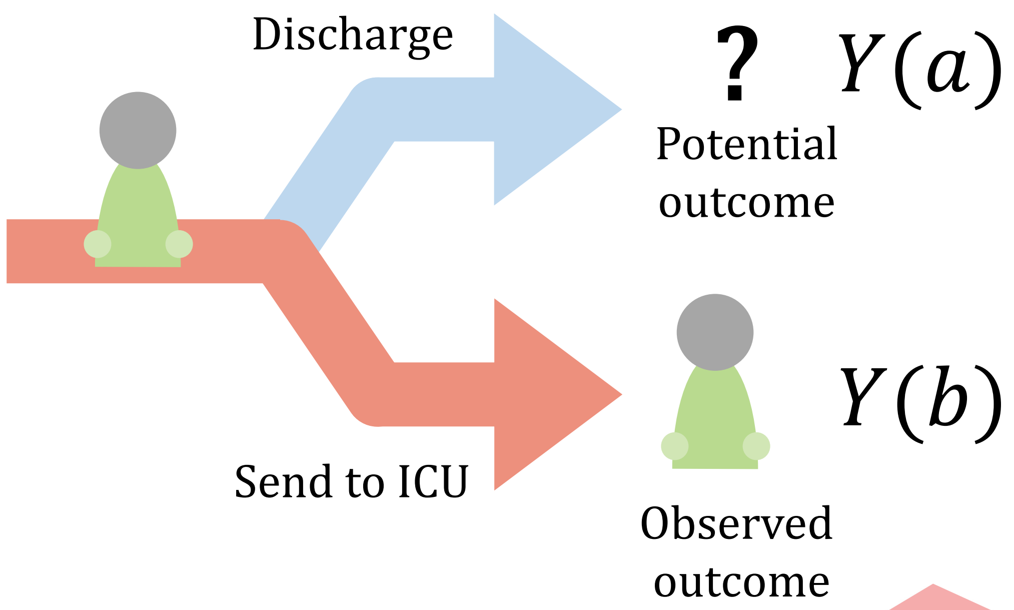
Average Potential Outcome: $E[Y(a)]$

Average Treatment Effect: $E[Y(a) - Y(b)]$

Average Potential Outcome: $E[Y(b)]$

I will refer to all these quantities as "causal effects", though that is more commonly used to refer to a contrast like the Average Treatment Effect

The challenge: We don't observe all outcomes!



Challenge: $E[Y(b)] \neq E[Y \mid T = b]$

Illustrative Example

- Suppose COVID-19 patients who are sent to the ICU have higher mortality rates
- Should we stop sending patients to the ICU?

Correlation does not imply causation

Correlation ~~does not~~ imply causation

Carefully chosen correlation can imply causation

(under some untestable assumptions)

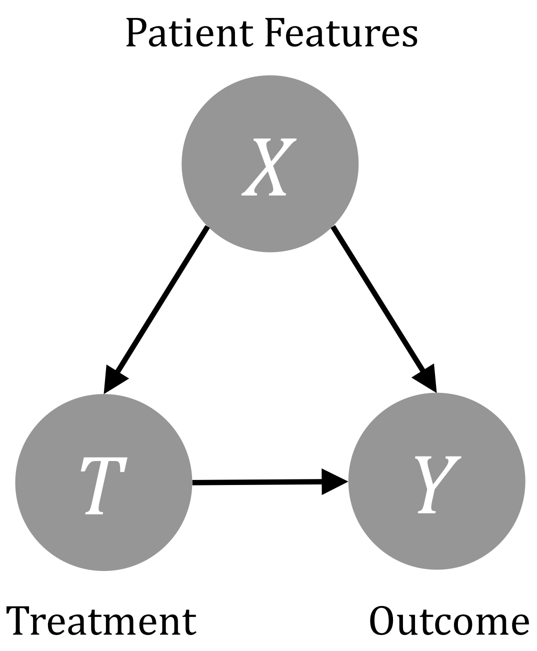
From Correlation to Causation

1 **Consistency** — When treatment t is given, we observe $Y(t)$

2 **No interference** — Your outcome is not impacted by the treatments of others

3 **No unmeasured confounding** — We have measured all relevant confounders

4 **Overlap / Coverage** — All types of patients receive all treatments*



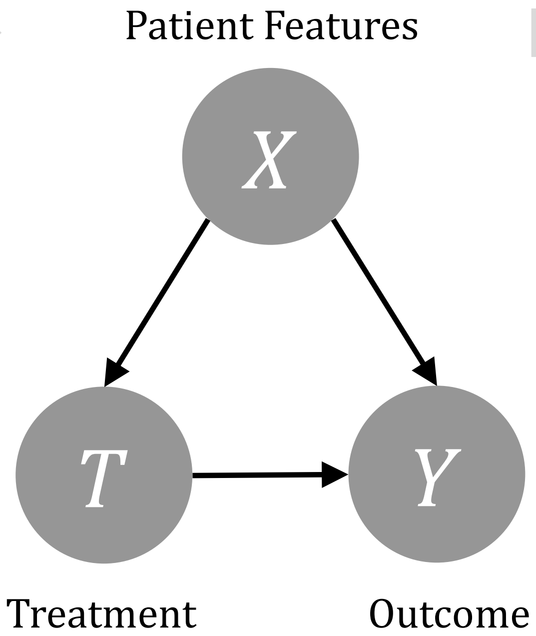
*More on this later, when we discuss policy evaluation

From Correlation to Causation

Causal Query

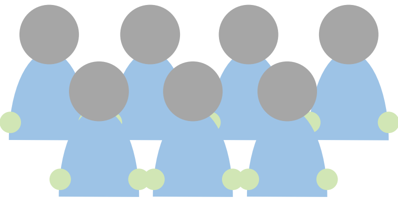
$$E[Y(t)]$$

Causal Assumptions



Statistical Query

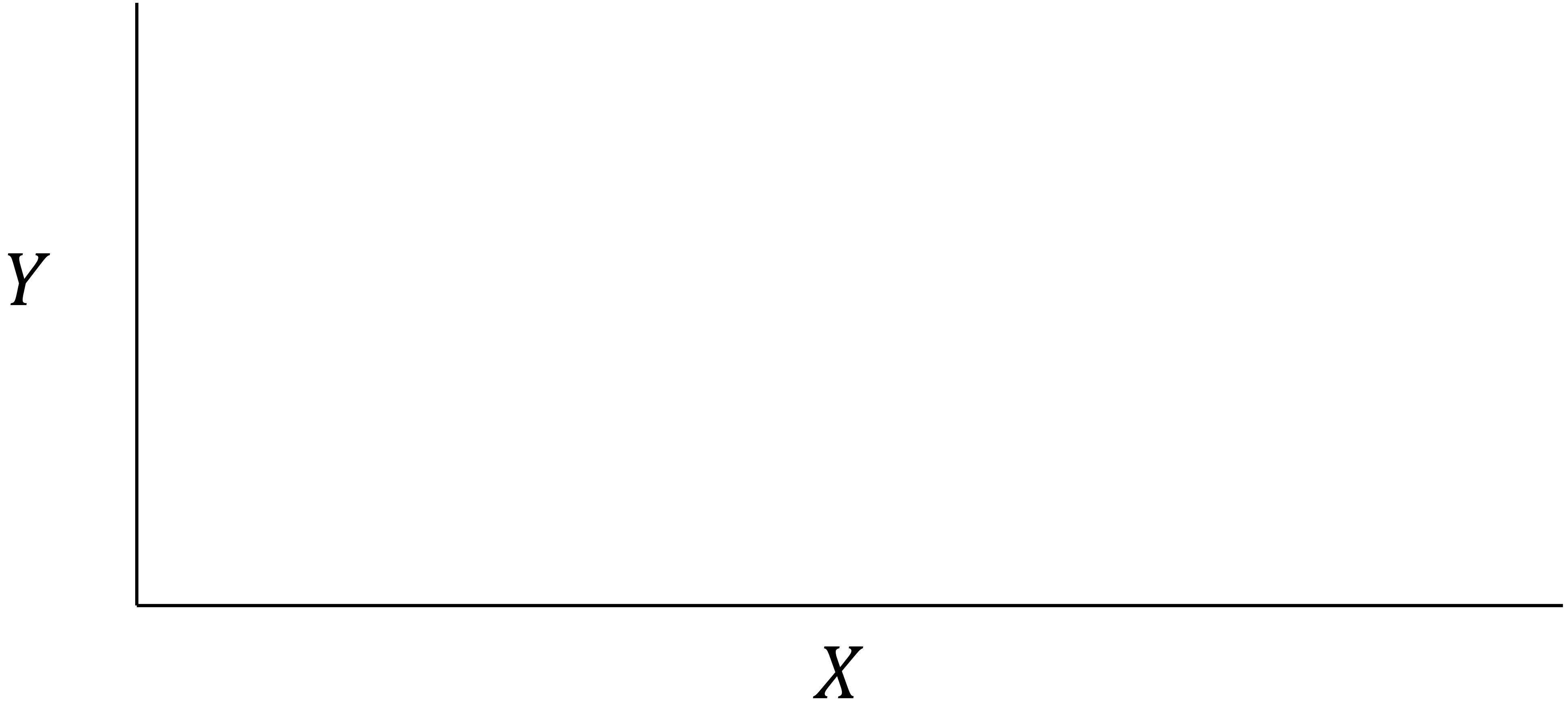
$$E\big[E[\,Y \mid X, T = t\,]\big]$$



Observational Data

Average Potential Outcomes

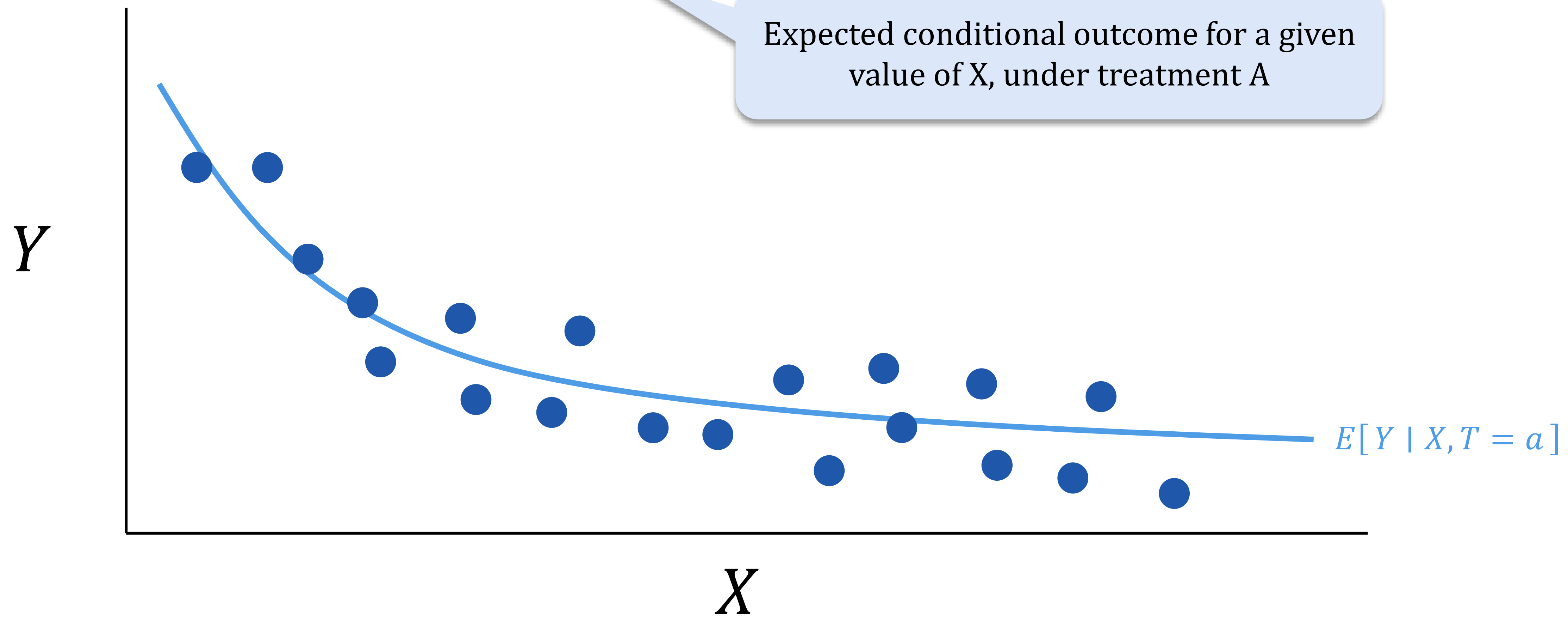
$$\mathrm{E}[Y(a)] = E_x\left[E[\,Y \mid X, T = a\,]\right]$$



Average Potential Outcomes

$$E[Y(a)] = E_x\big[E[\,Y \mid X, T = a\,]\big]$$

Expected conditional outcome for a given value of X, under treatment A

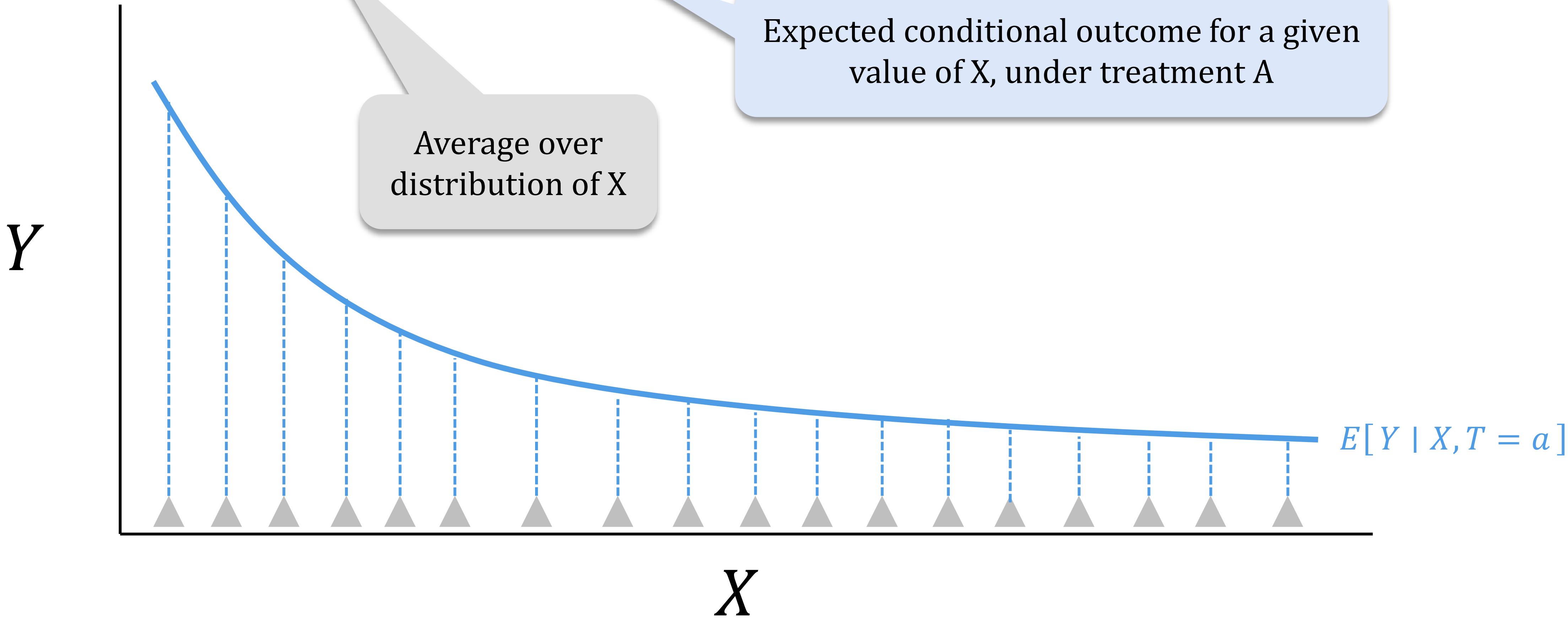


Average Potential Outcomes

$$E[Y(a)] = E_x\left[E[Y \mid X, T = a]\right]$$

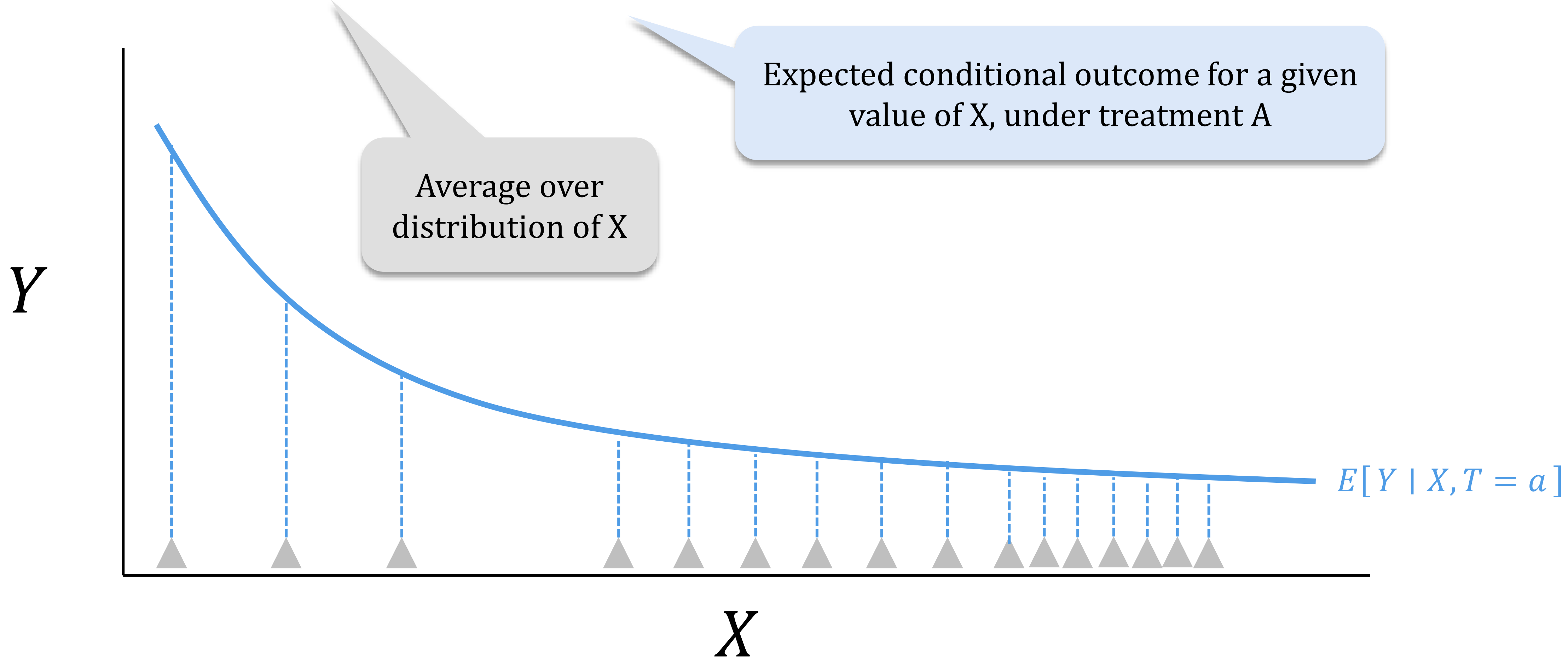
Average over distribution of X

Expected conditional outcome for a given value of X, under treatment A



Average Potential Outcomes

$$\mathrm{E}[Y(a)] = E_x\left[E[Y \mid X, T = a]\right]$$

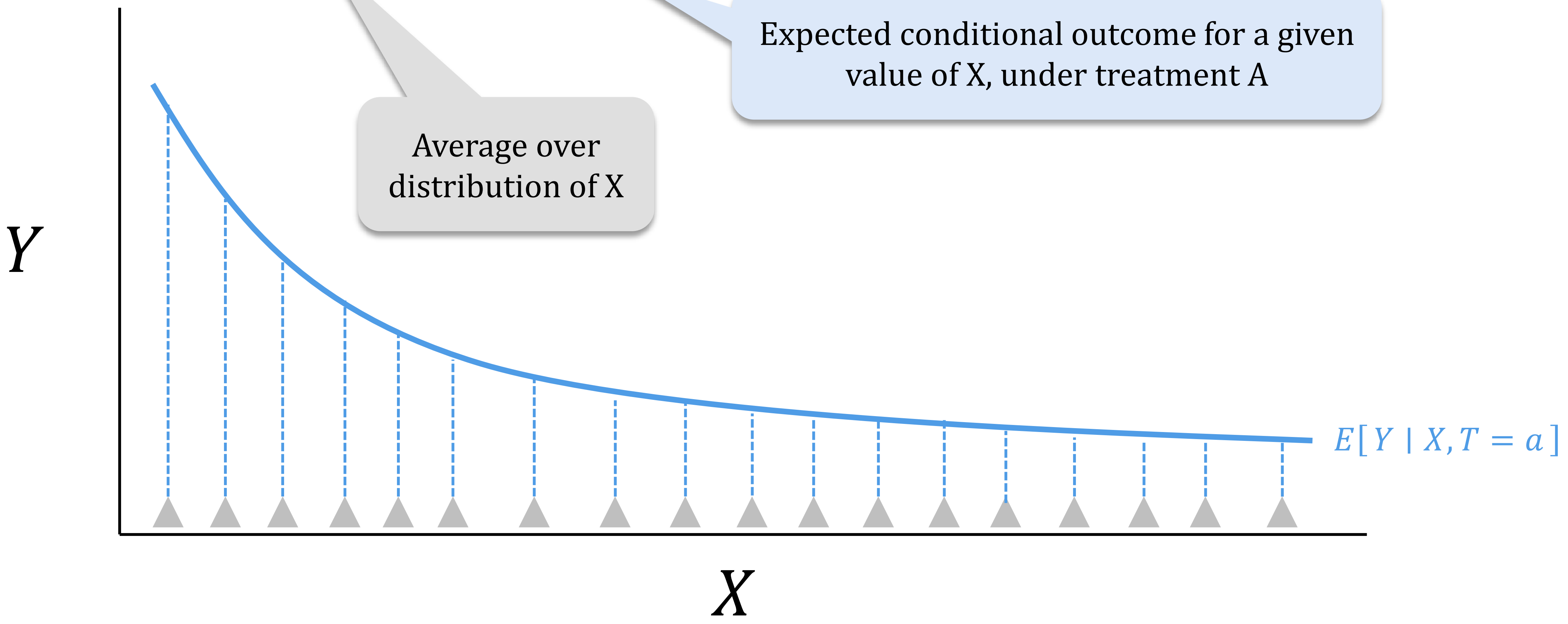


Average Potential Outcomes

$$E[Y(a)] = E_x\left[E[Y \mid X, T = a]\right]$$

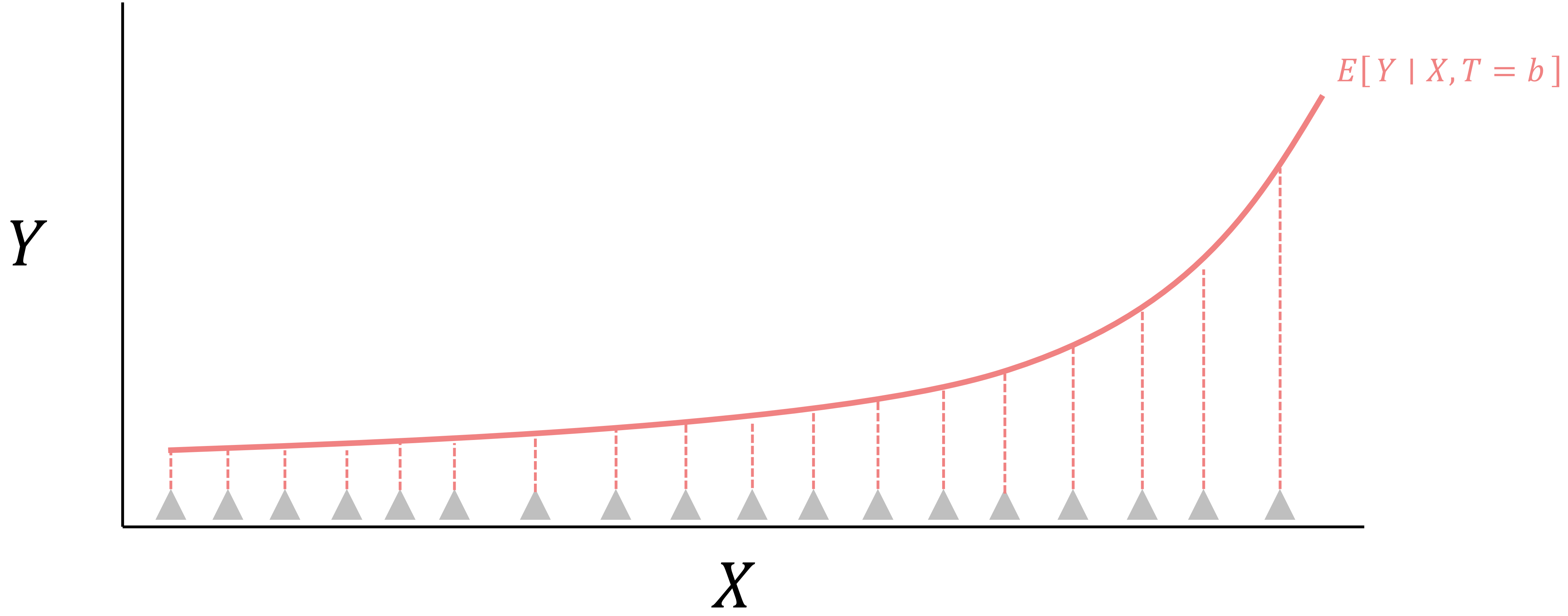
Average over distribution of X

Expected conditional outcome for a given value of X, under treatment A

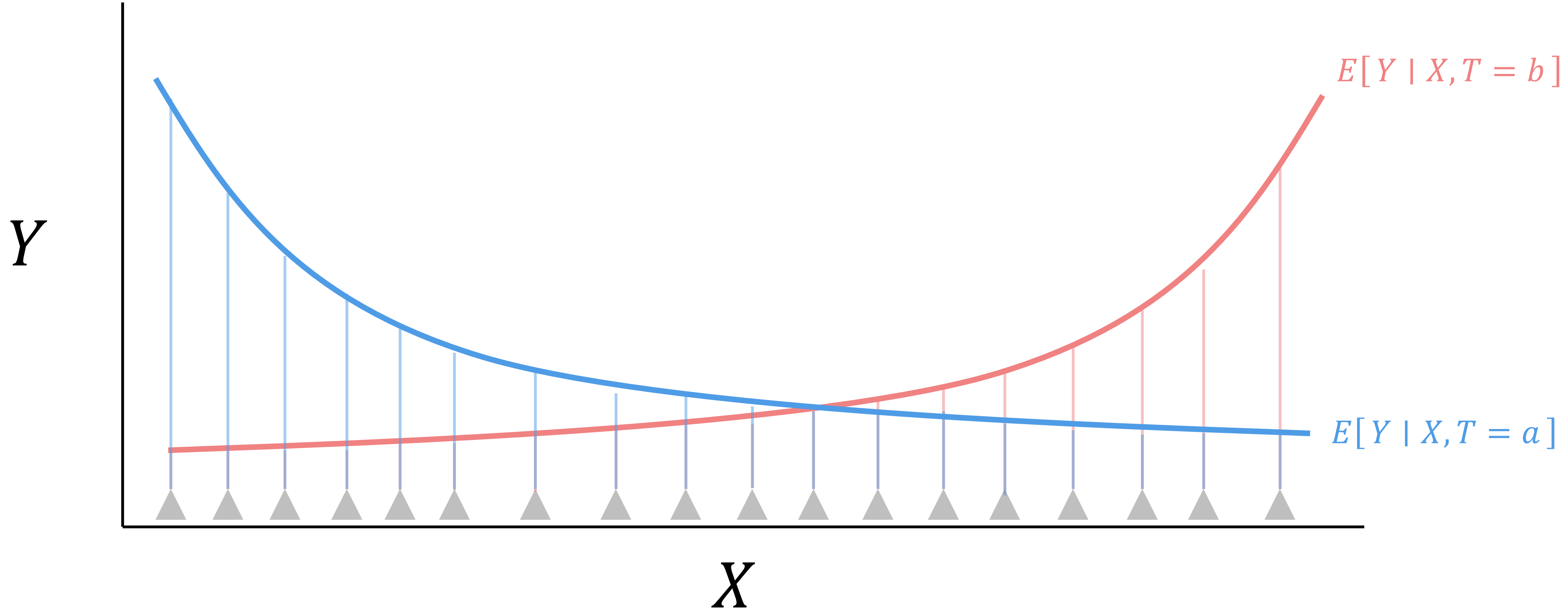


What about a different treatment?

$$\mathrm{E}[Y(b)] = E_x\left[E[Y \mid X, T = b]\right]$$

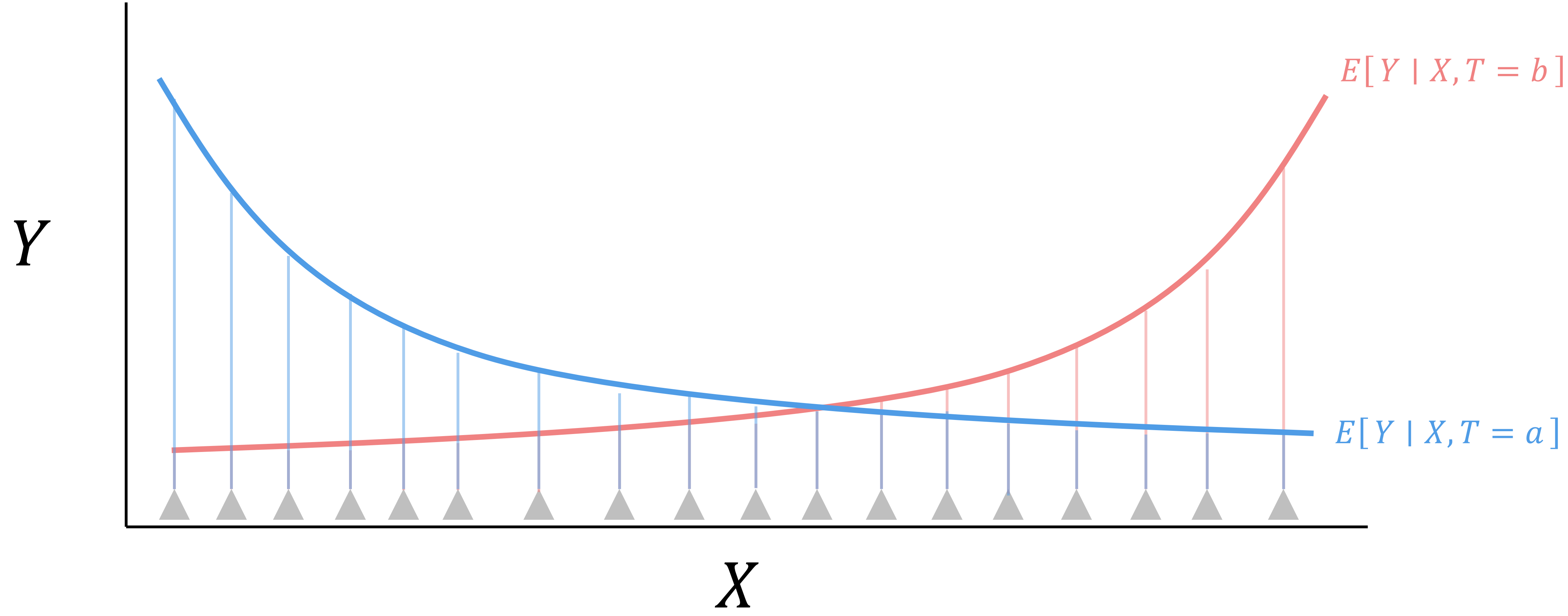


What is the Average Treatment Effect?



What is the Average Treatment Effect?

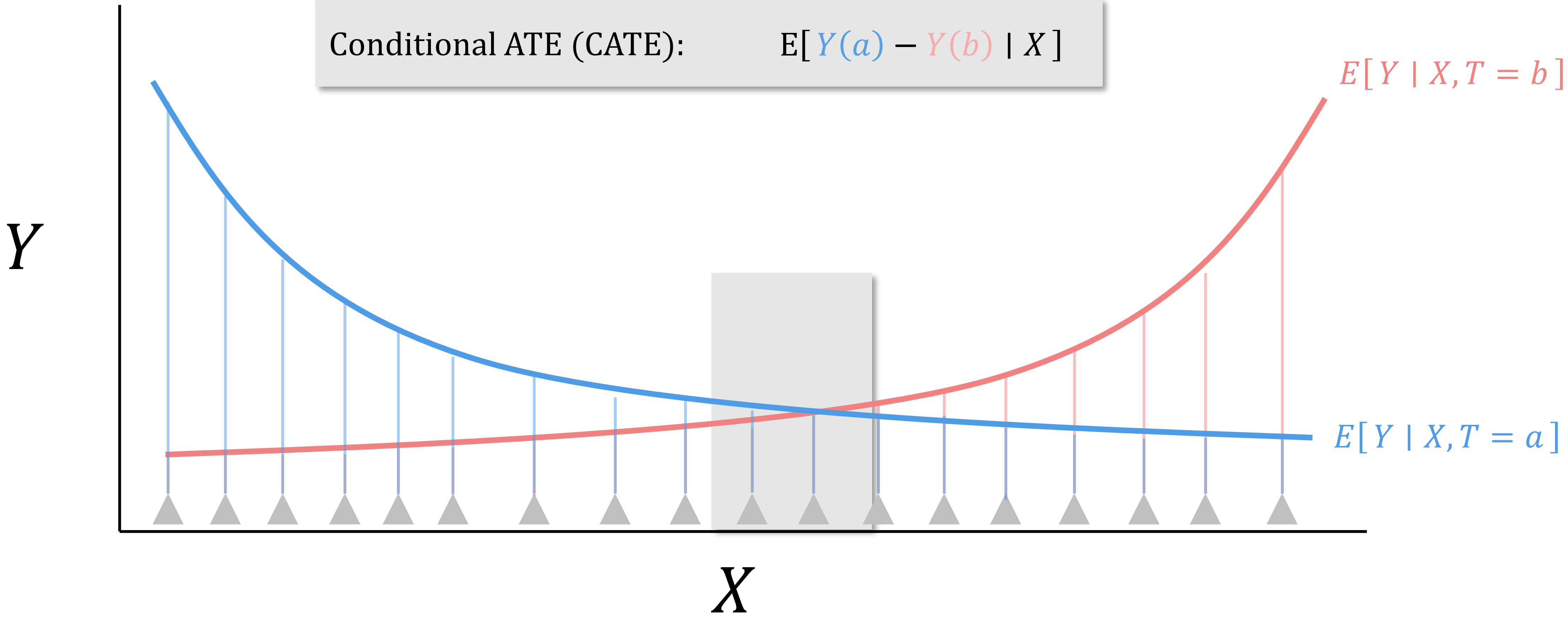
Average Treatment Effect (ATE): $\mathrm{E}[Y(a) - Y(b)]$



What is the (Conditional) Average Treatment Effect?

Average Treatment Effect (ATE): $\mathrm{E}[Y(a) - Y(b)]$

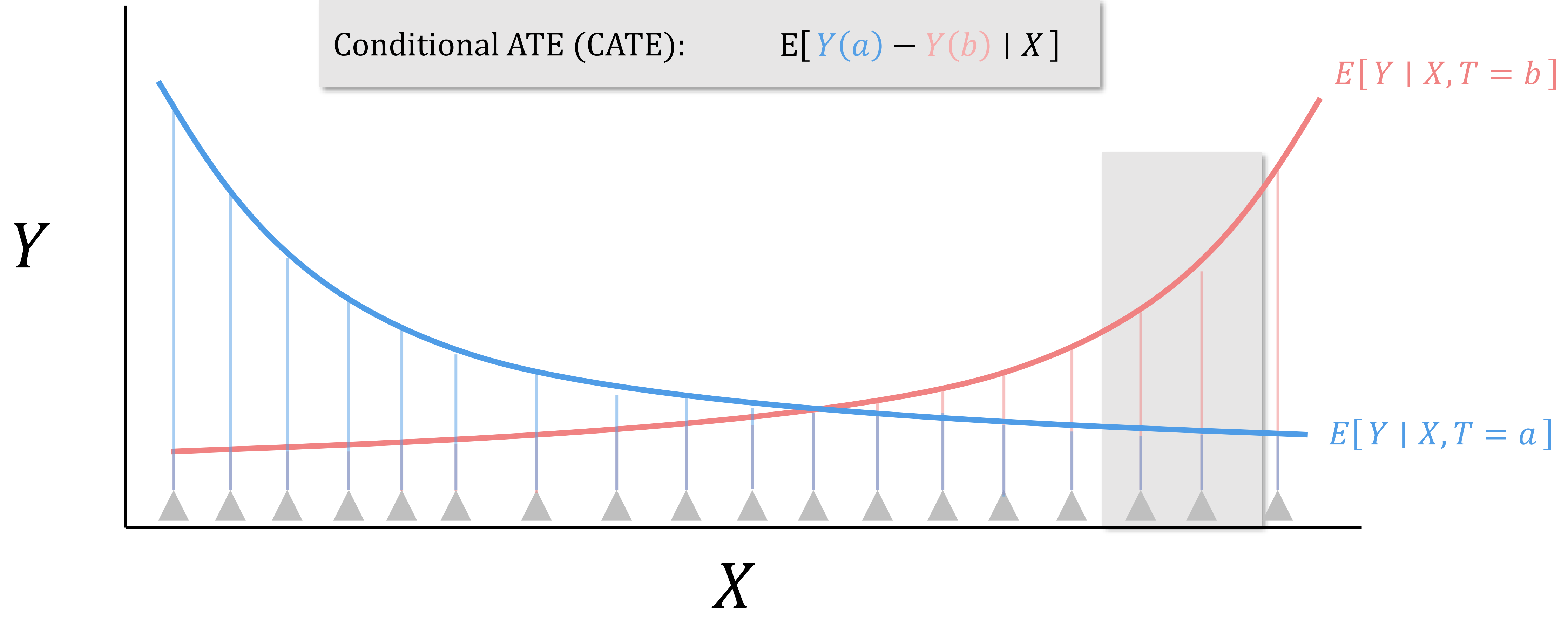
Conditional ATE (CATE): $\mathrm{E}[Y(a) - Y(b) \mid X]$



What is the (Conditional) Average Treatment Effect?

Average Treatment Effect (ATE): $\mathrm{E}[Y(a) - Y(b)]$

Conditional ATE (CATE): $\mathrm{E}[\, Y(a) - Y(b) \mid X\,]$



Goals

- By the end of this primer, you should understand…
  - What a **causal effect** is, and when we can infer it from observational data

    "What is the effect of treatment vs. placebo?"
  - How **causal effect estimation** is a special case of **policy evaluation**

    "What would happen if I implemented a new set of guidelines?"
  - How to use policy evaluation to **learn optimal policies**

    "What is the optimal treatment strategy?"

Policy Evaluation

What is a policy?

- A policy maps from **observed features** to **recommended actions**

$$\pi : X \to A$$

x → a

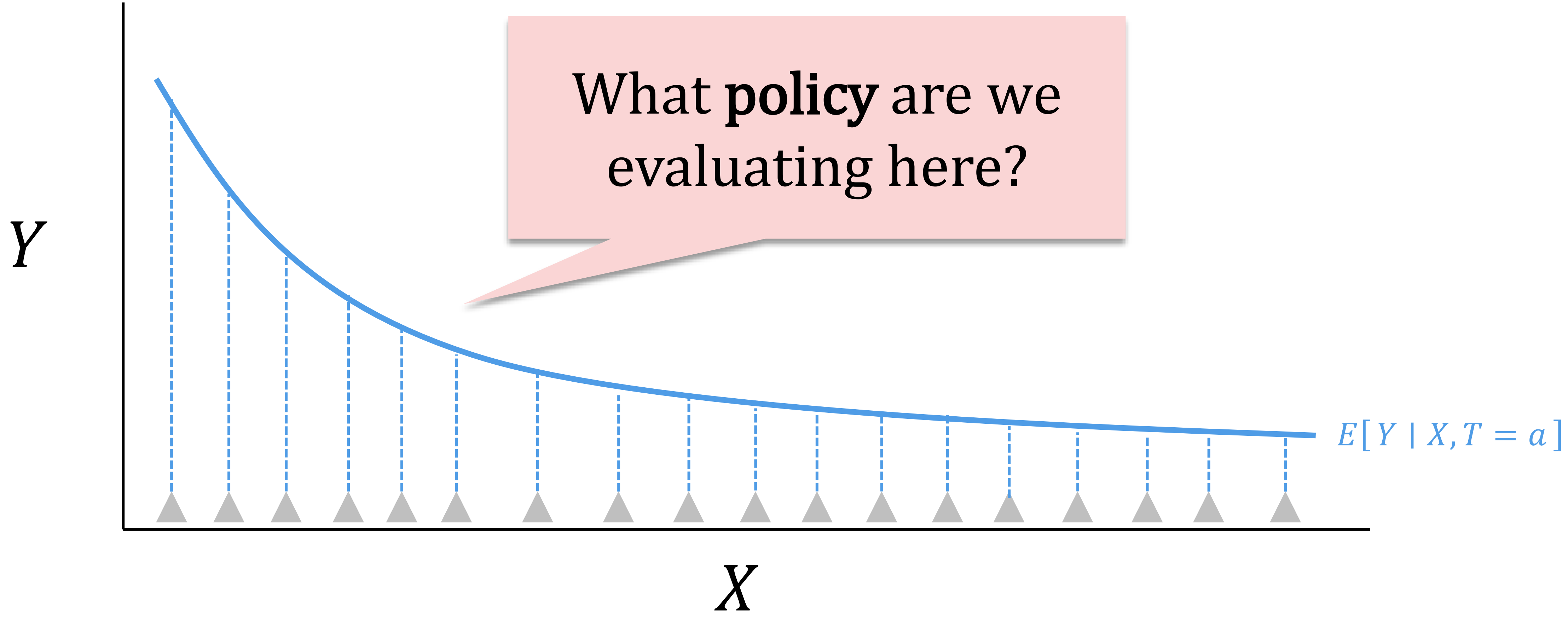
Patient Features

Action

(Or a distribution over actions)

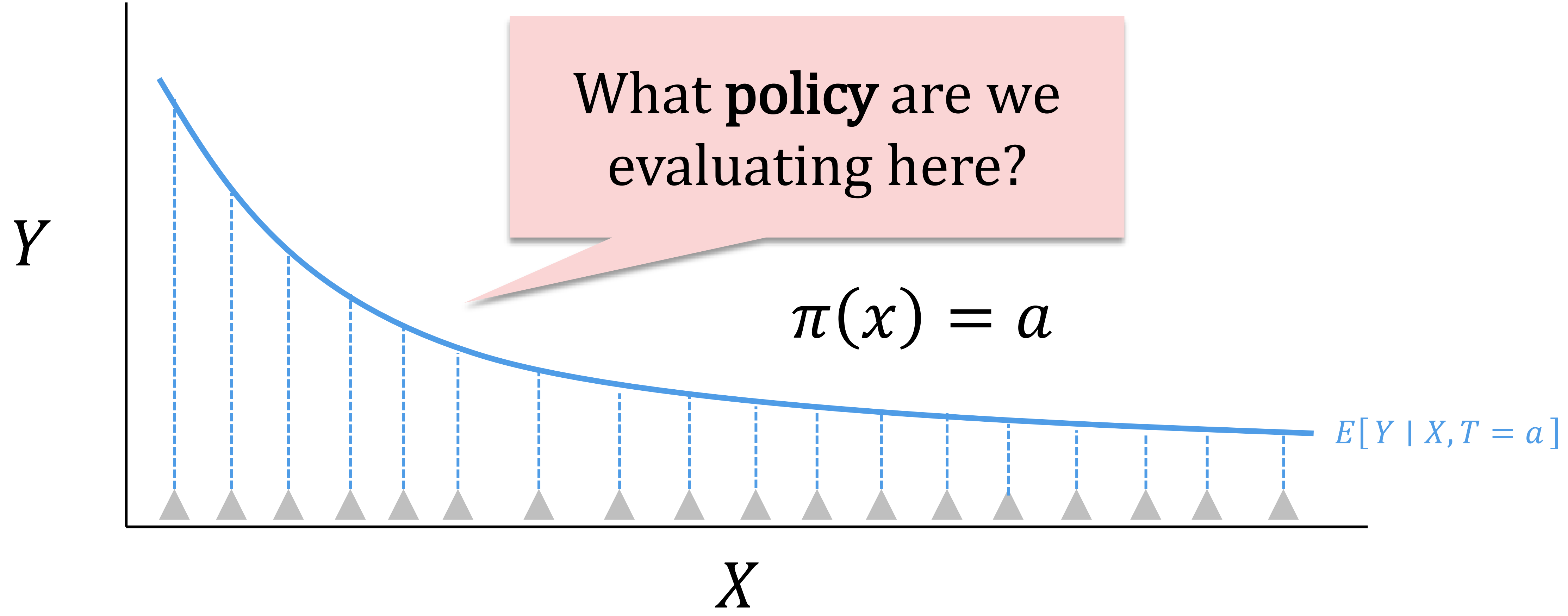
Causal Effects as Policy Evaluation

$$E[Y(a)] = E_x[E[Y \mid X, T = a]]$$



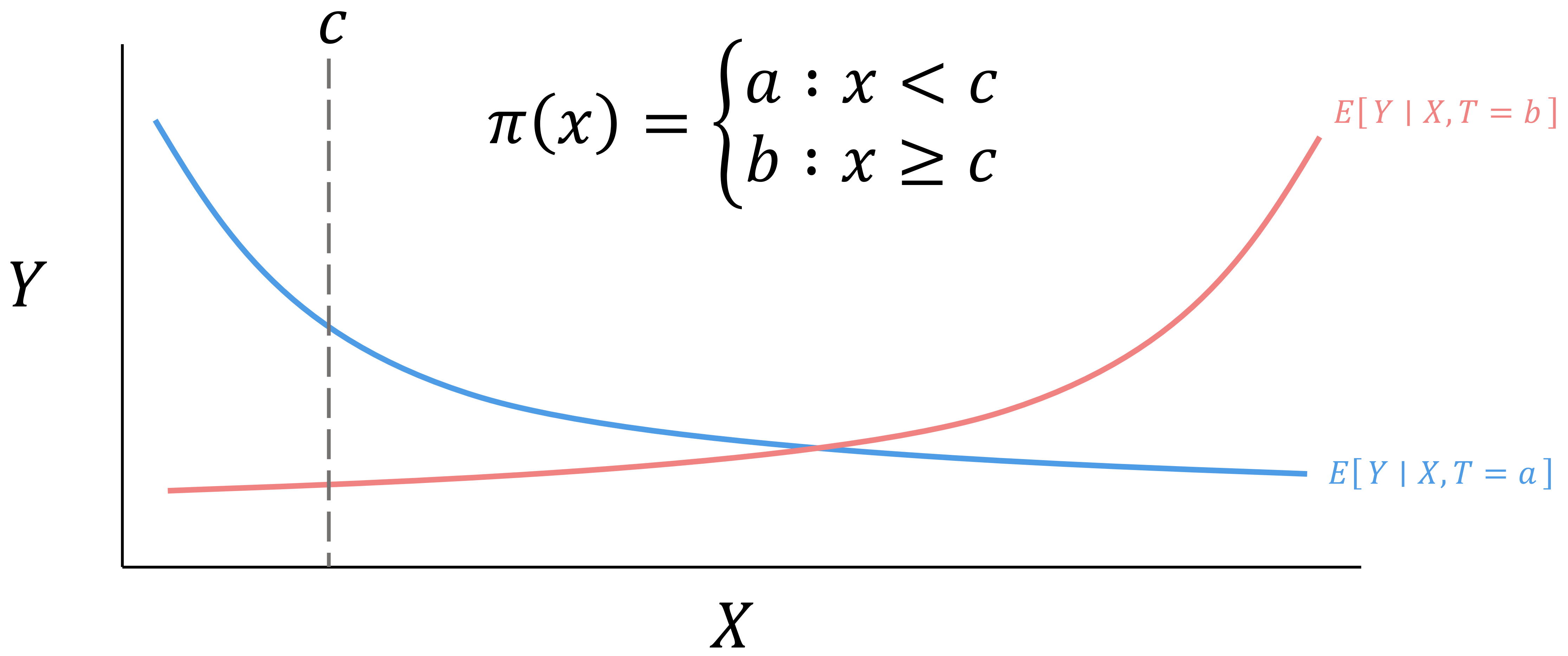
Causal Effects as Policy Evaluation

$$\mathrm{E}[Y(a)] = E_x\left[E[\,Y \mid X, T = a\,]\right]$$



What is the value of a policy?

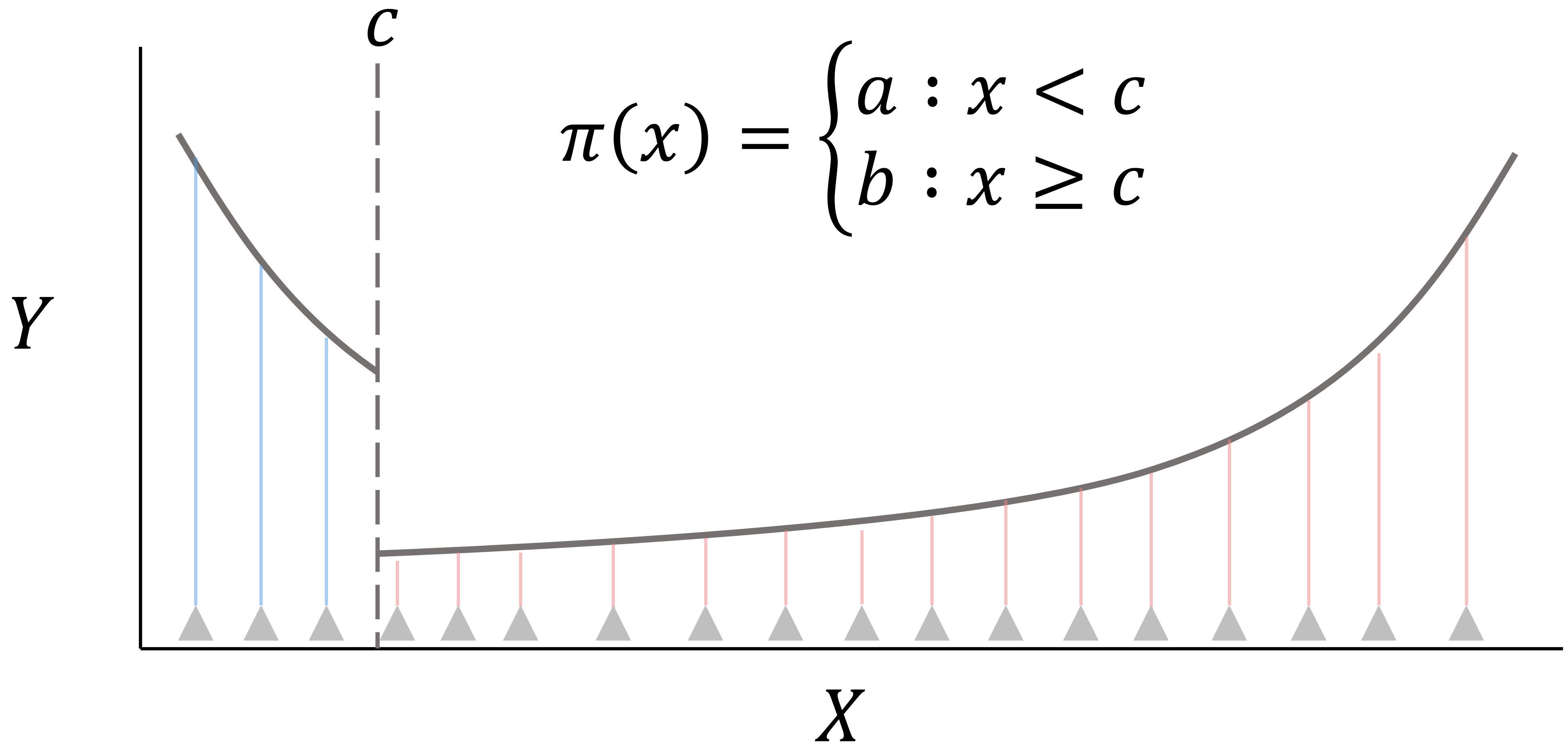
$$E[Y(\pi(x))] = E_x[E[Y \mid X, T = \pi(x)]]$$



What is the value of a policy?

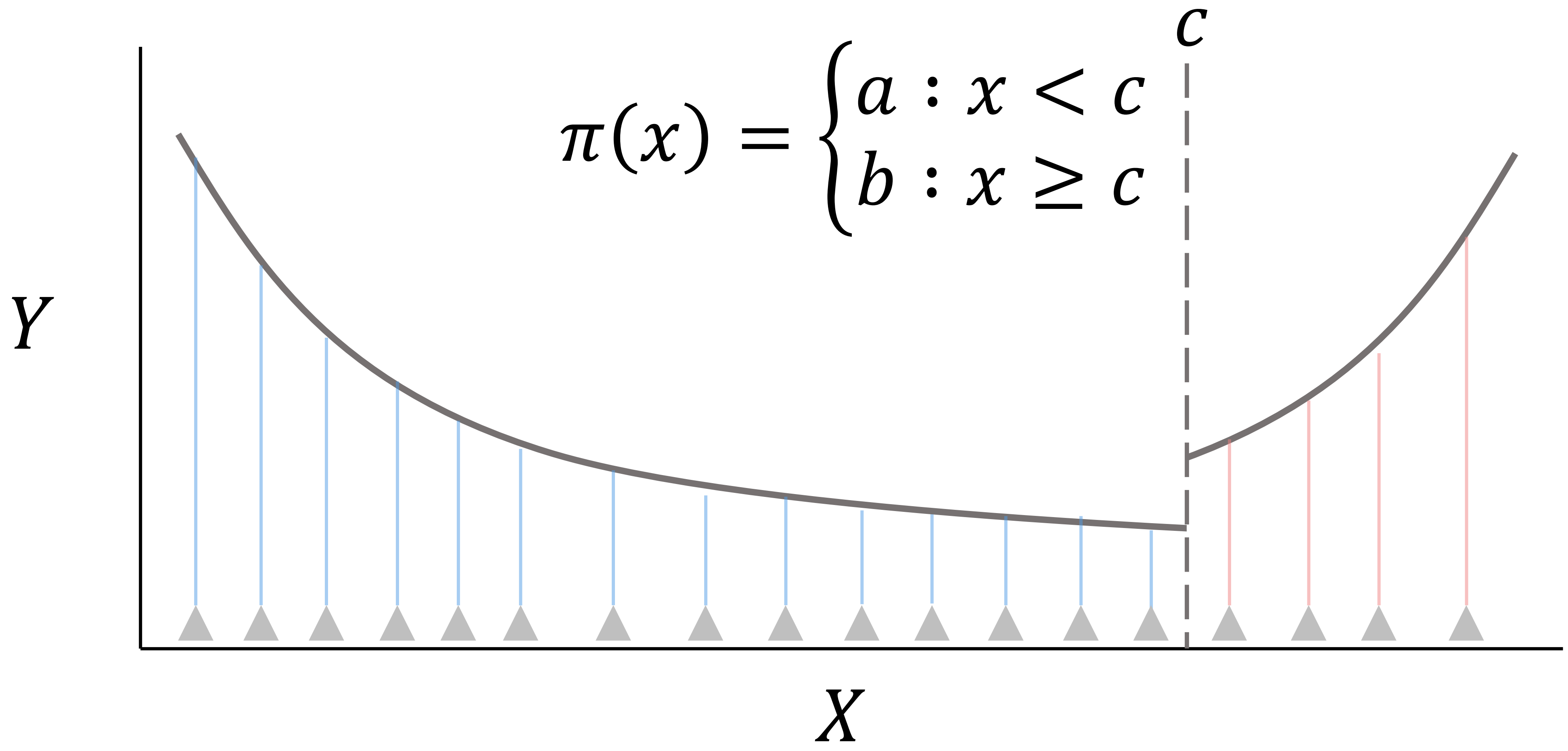
$$\mathrm{E}\left[Y\left(\pi(x)\right)\right] = E_x\left[E[\,Y \mid X, T = \pi(x)\,]\right]$$

$$\pi(x) = \begin{cases} a : x < c \\ b : x \geq c \end{cases}$$



What is the value of a policy?

$$\mathrm{E}\left[Y\left(\pi(x)\right)\right] = E_x\left[E[\,Y \mid X, T = \pi(x)\,]\right]$$



How do we evaluate a policy?

$$\mathrm{E}\left[Y\left(\pi(x)\right)\right] = E_x\left[E[\,Y \mid X, T = \pi(x)\,]\right]$$

1

Outcome Regression

$$\hat{E}\left[Y\left(\pi(x)\right)\right] = \frac{1}{n}\sum_{i=1}^{n} f\left(x_i, \pi(x_i)\right)$$

$$f(x,t) \approx E[Y \mid X = x, T = t]$$

How do we evaluate a policy?

$$\mathrm{E}\left[Y\left(\pi(x)\right)\right] = E_x\left[E[\,Y \mid X, T = \pi(x)\,]\right]$$

1 **Outcome Regression**

$$\hat{E}\left[Y\left(\pi(x)\right)\right] = \frac{1}{n}\sum_{i=1}^{n} f\left(x_i, \pi(x_i)\right)$$

$$f(x,t) \approx E[Y \mid X = x, T = t]$$

2 **Propensity Reweighting**

$$\hat{E}\left[Y\left(\pi(x)\right)\right] = \frac{1}{n}\sum_{i=1}^{n} \frac{1\{t_i = \pi(x_i)\}}{e(x_i, t_i)} Y_i$$

$$e(x,t) \approx p[T = t \mid X = x]$$

How do we evaluate a policy?

$$\mathrm{E}\left[Y\left(\pi(x)\right)\right] = E_x\left[E[\,Y \mid X, T = \pi(x)\,]\right]$$

1 **Outcome Regression**

$$\hat{E}\left[Y\left(\pi(x)\right)\right] = \frac{1}{n}\sum_{i=1}^{n} f\left(x_i, \pi(x_i)\right) \qquad f(x,t) \approx E[Y \mid X = x, T = t]$$

2 **Propensity Reweighting**

$$\hat{E}\left[Y\left(\pi(x)\right)\right] = \frac{1}{n}\sum_{i=1}^{n} \frac{1\{t_i = \pi(x_i)\}}{e(x_i, t_i)} Y_i \qquad e(x,t) \approx p[T = t \mid X = x]$$

3 **“Doubly Robust”**

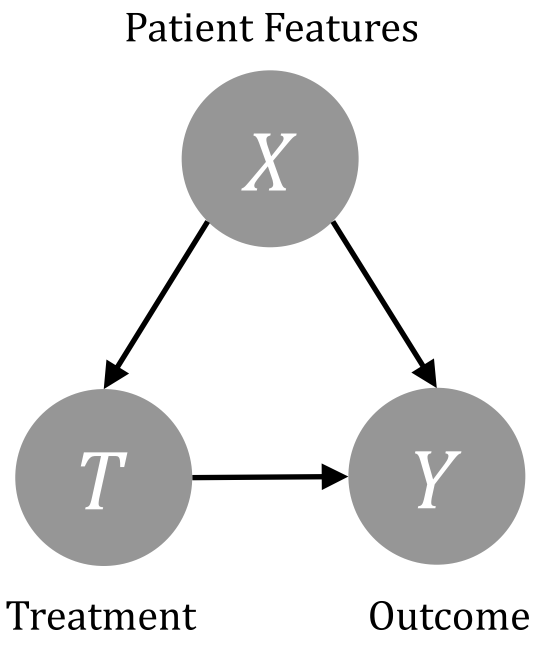
$$\hat{E}\left[Y\left(\pi(x)\right)\right] = \frac{1}{n}\sum_{i=1}^{n} \frac{1\{t_i = \pi(x_i)\}}{e(x_i, t_i)} \left(Y_i - f\left(x_i, \pi(x_i)\right)\right) + f\left(x_i, \pi(x_i)\right)$$

Which policies can we evaluate?

(alternatively: For what types of patients can we evaluate a fixed policy?)

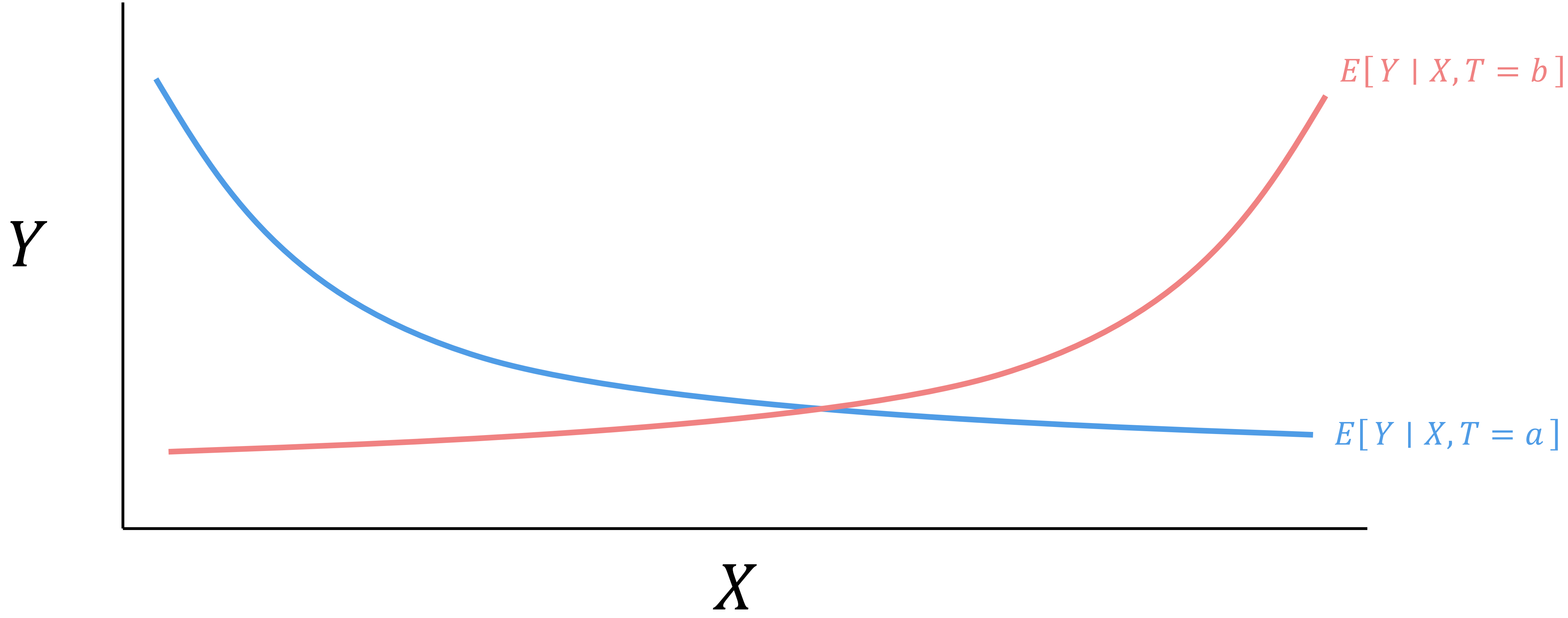
Which policies can we evaluate?

1. **Consistency** — When treatment t is given, we observe $Y(t)$
2. **No interference** — Your outcome is not impacted by the treatments of others
3. **No unmeasured confounding** — We have measured all relevant confounders
4. **Overlap / Coverage** — All types of patients receive all treatments*



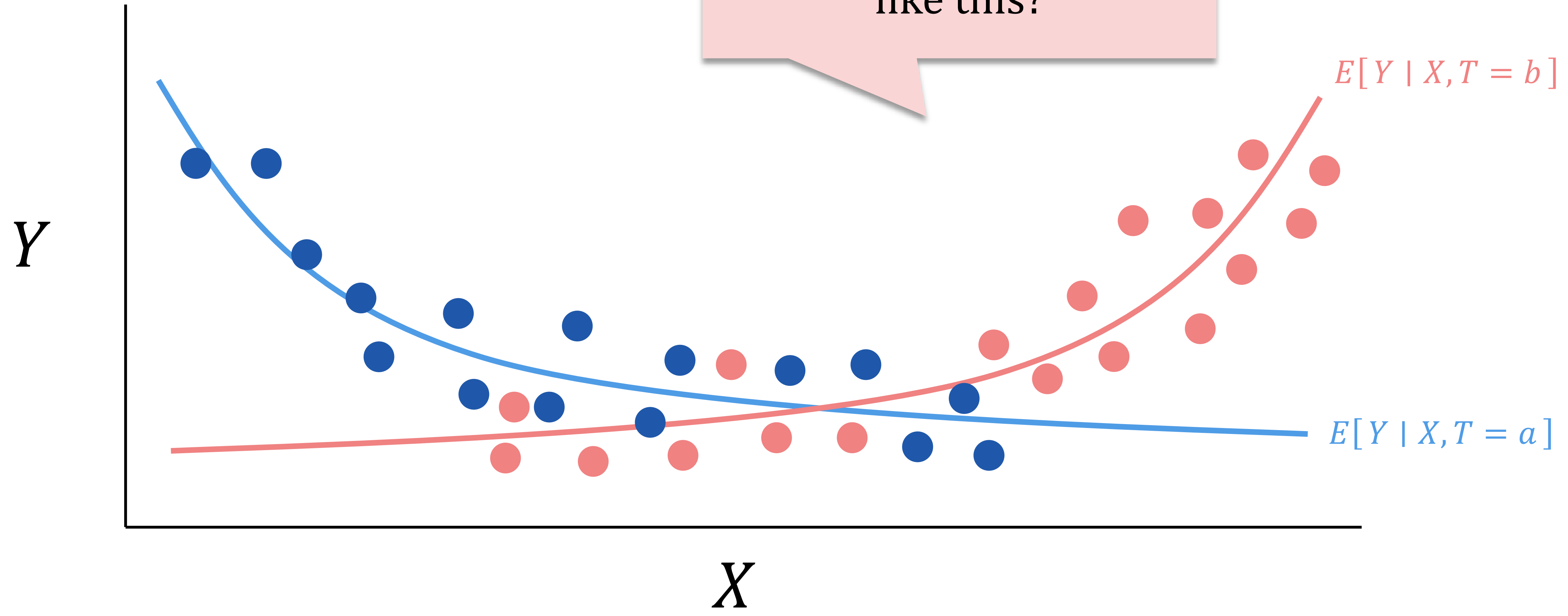
*More on this later, when we discuss policy evaluation

Which policies can we evaluate?

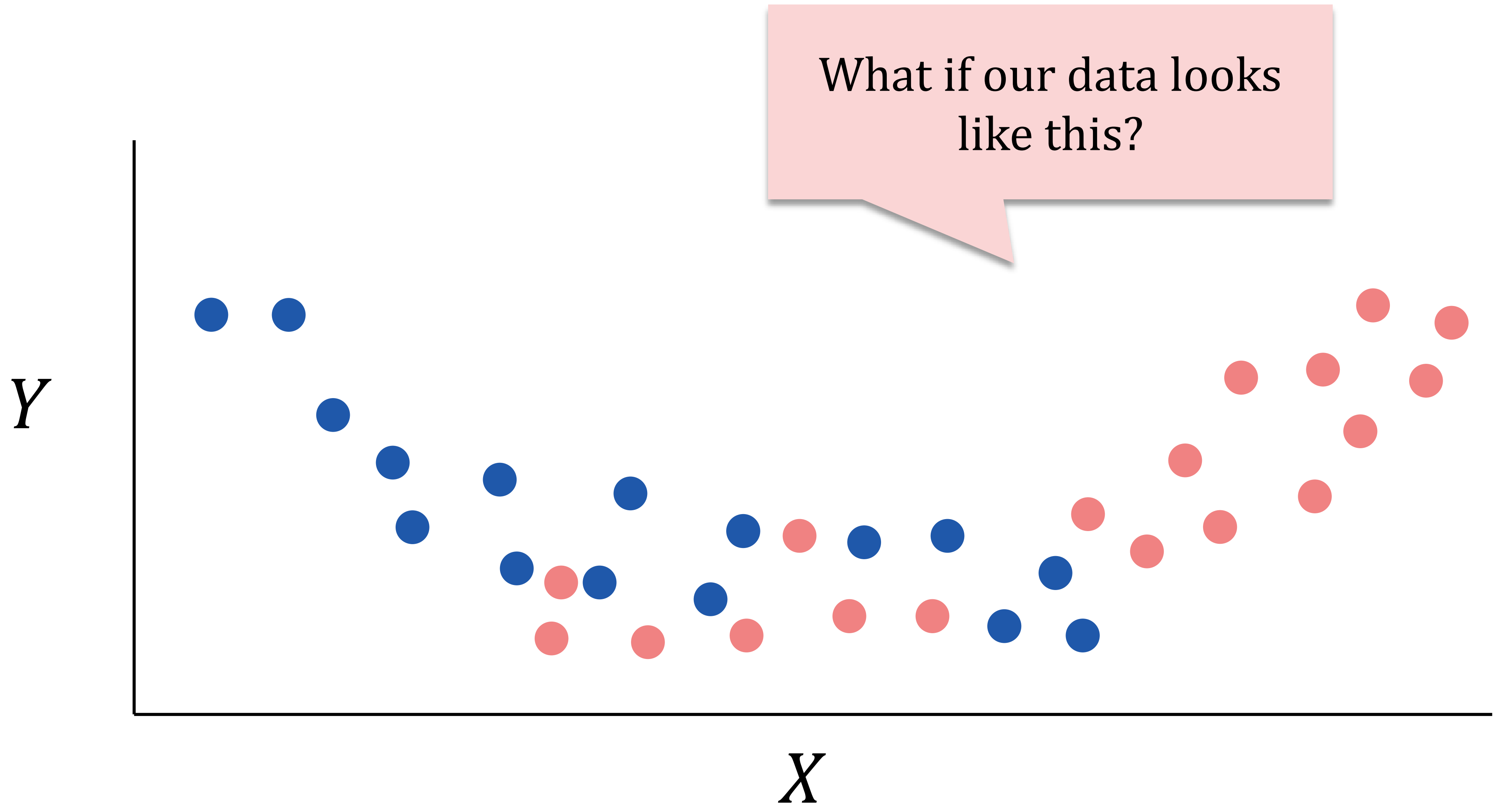


Which policies can we evaluate?

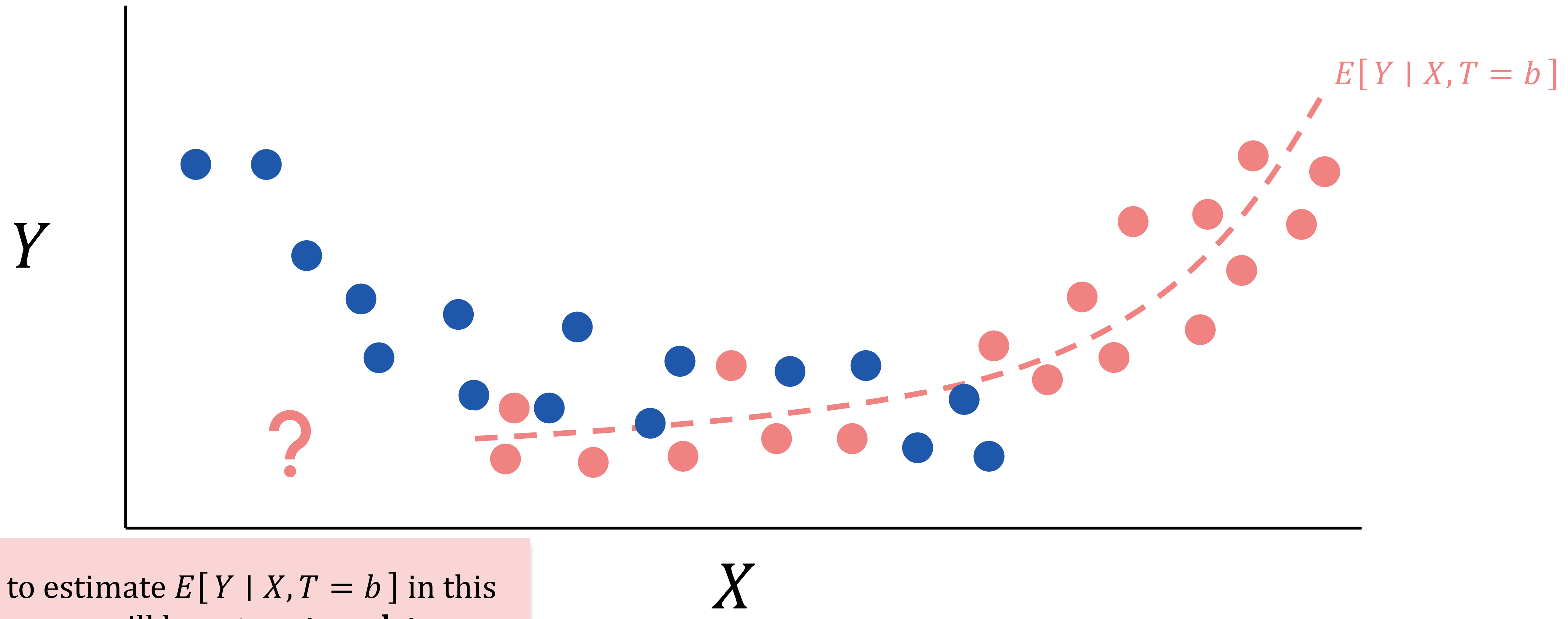
What if our data looks like this?



Which policies can we evaluate?

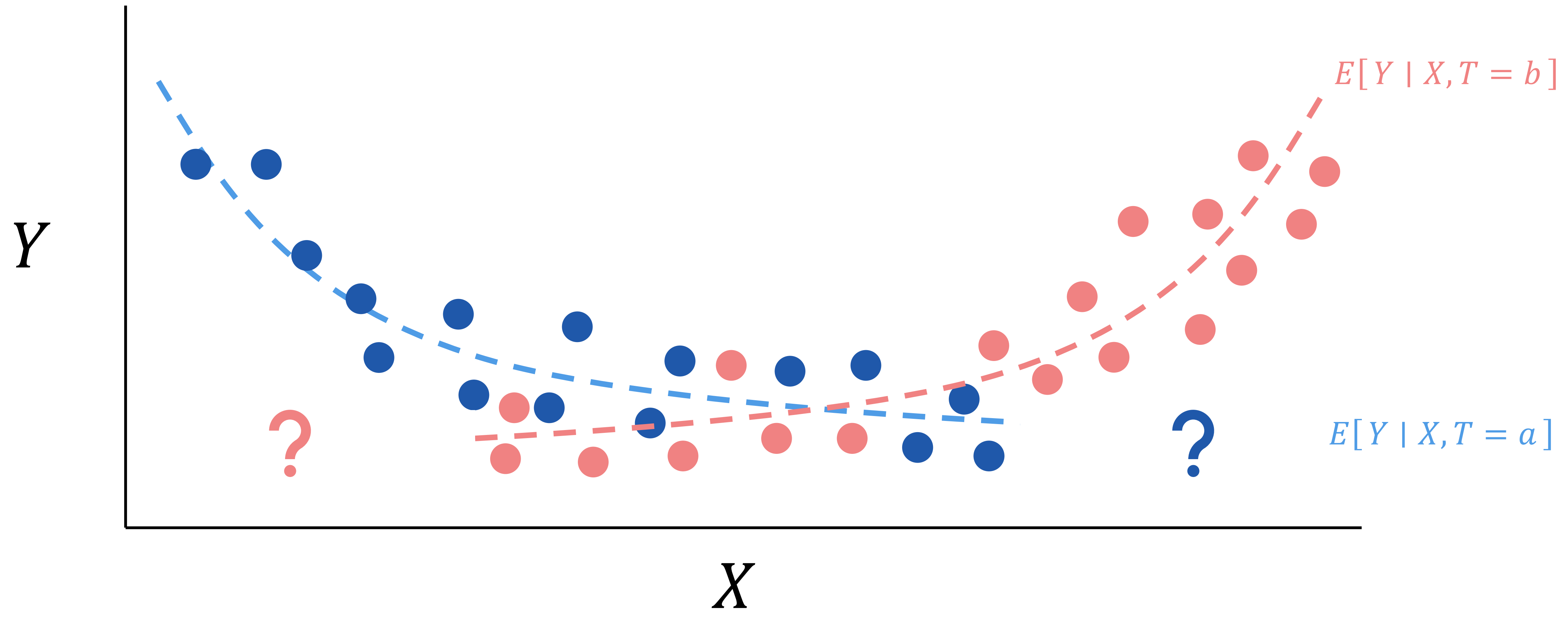


Which policies can we evaluate?



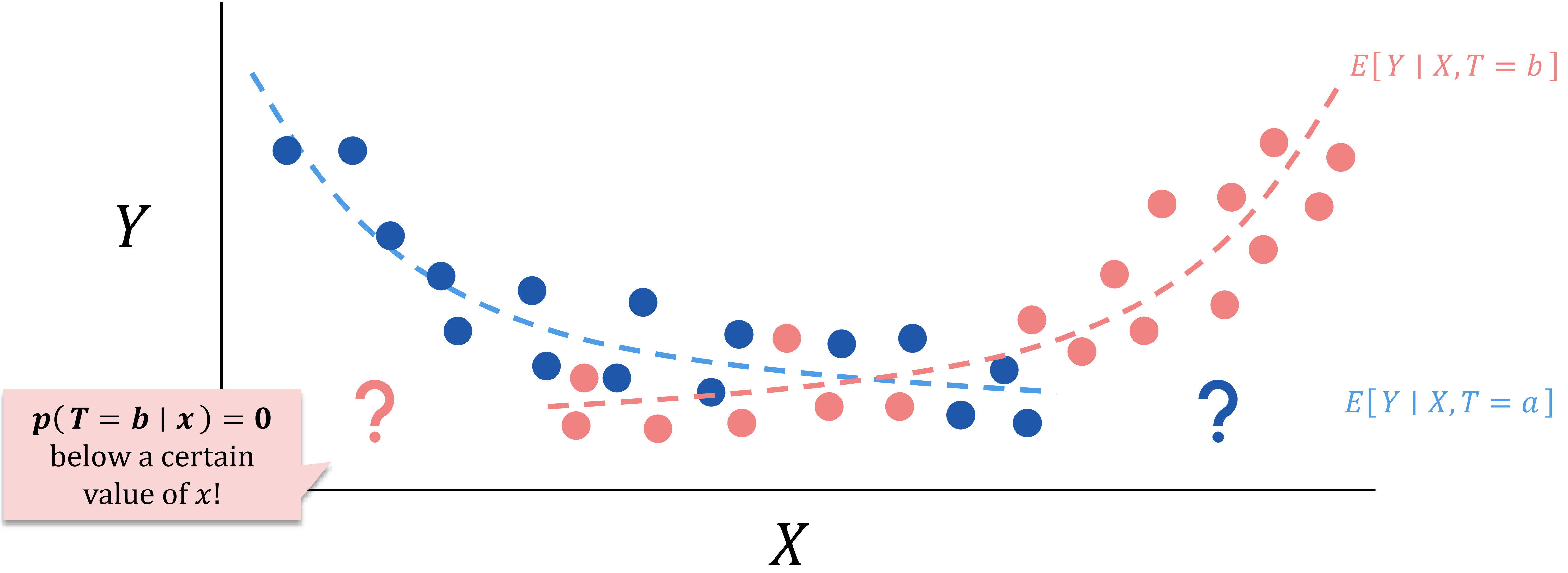
In order to estimate $E[Y \mid X, T = b]$ in this region, we will have to **extrapolate**

Which policies can we evaluate?



Which policies can we evaluate?

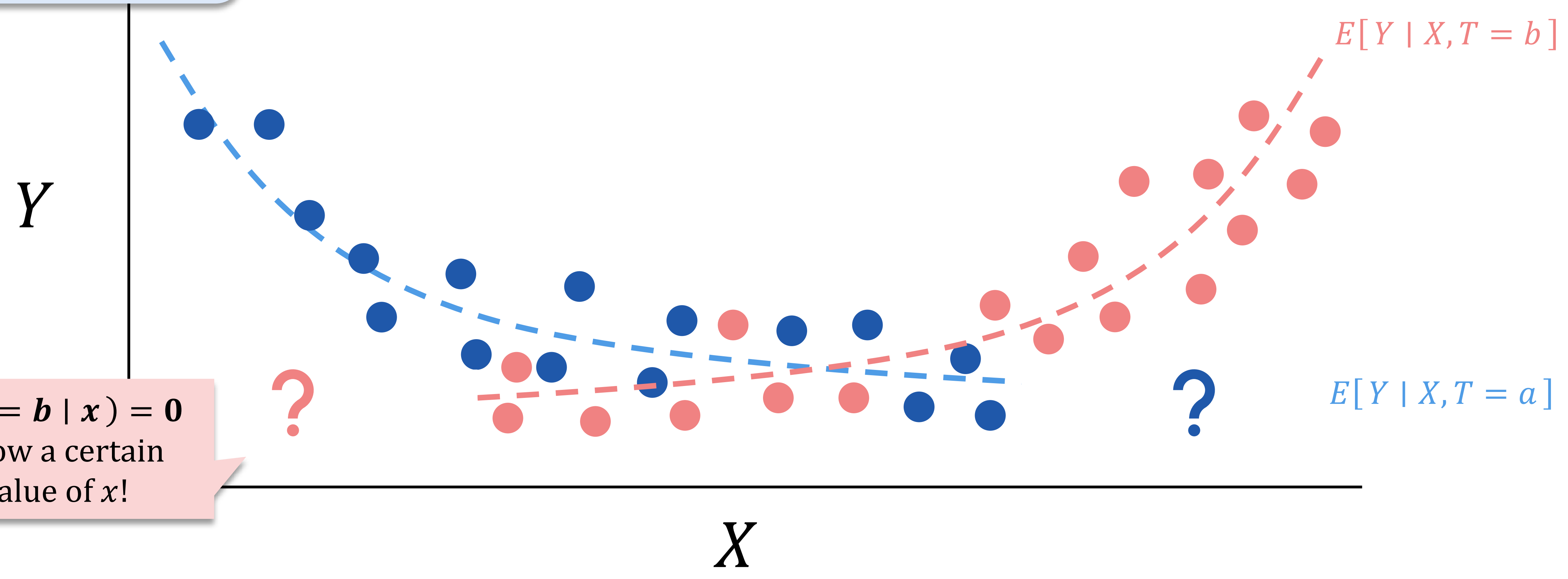
Coverage: For all x, $\pi(x) = t \Rightarrow p(t \mid x) > 0$



Which policies can we evaluate?

Depends on policy! Does not need to hold for all t

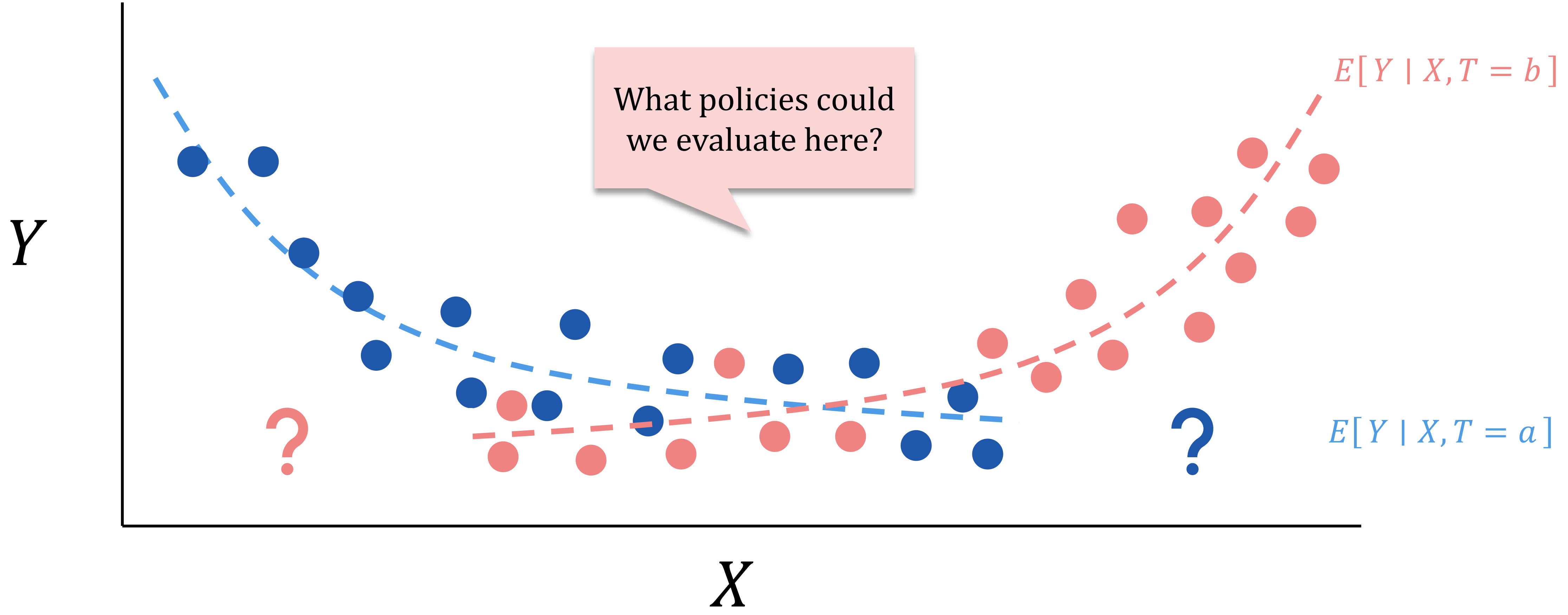
Coverage: For all x, $\pi(x) = t$ ⟶ $p(t \mid x) > 0$



$\boldsymbol{p(T = b \mid x) = 0}$ below a certain value of x!

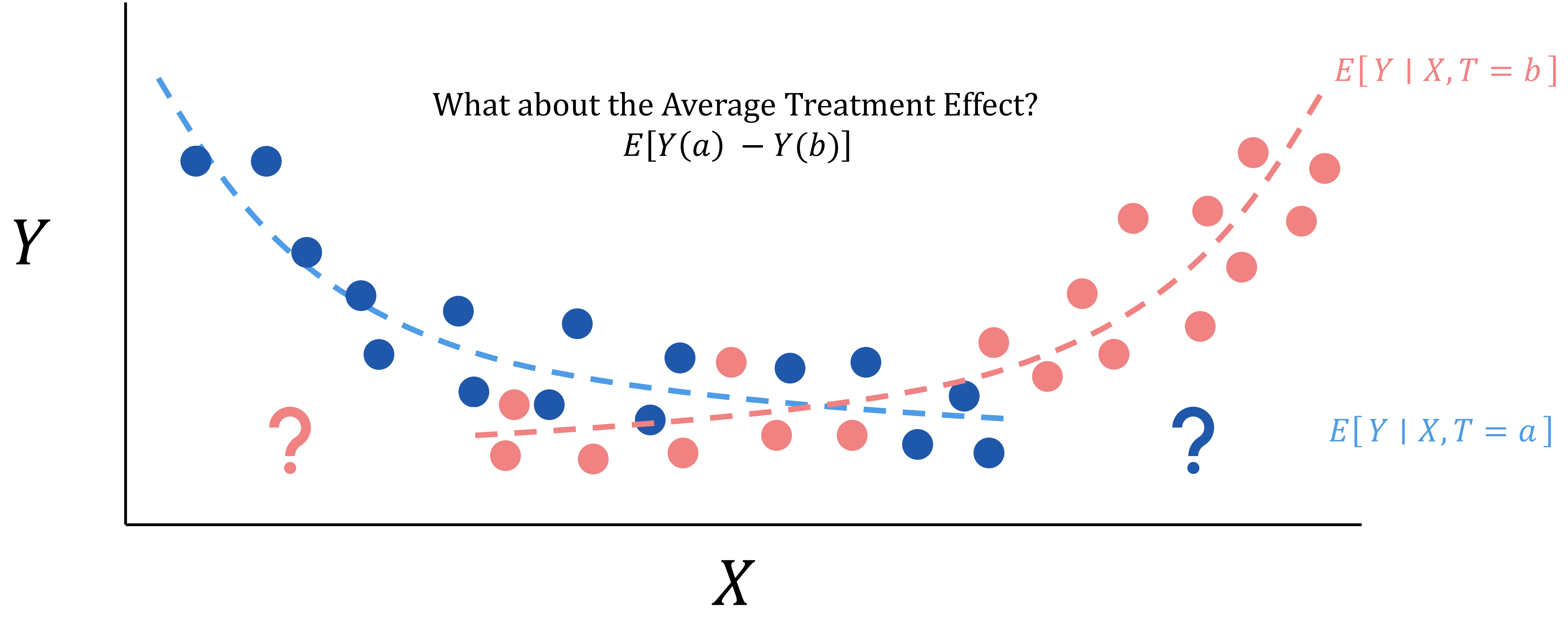
Which policies can we evaluate?

Coverage: For all x, $\pi(x) = t$ ⟶ $p(t \mid x) > 0$



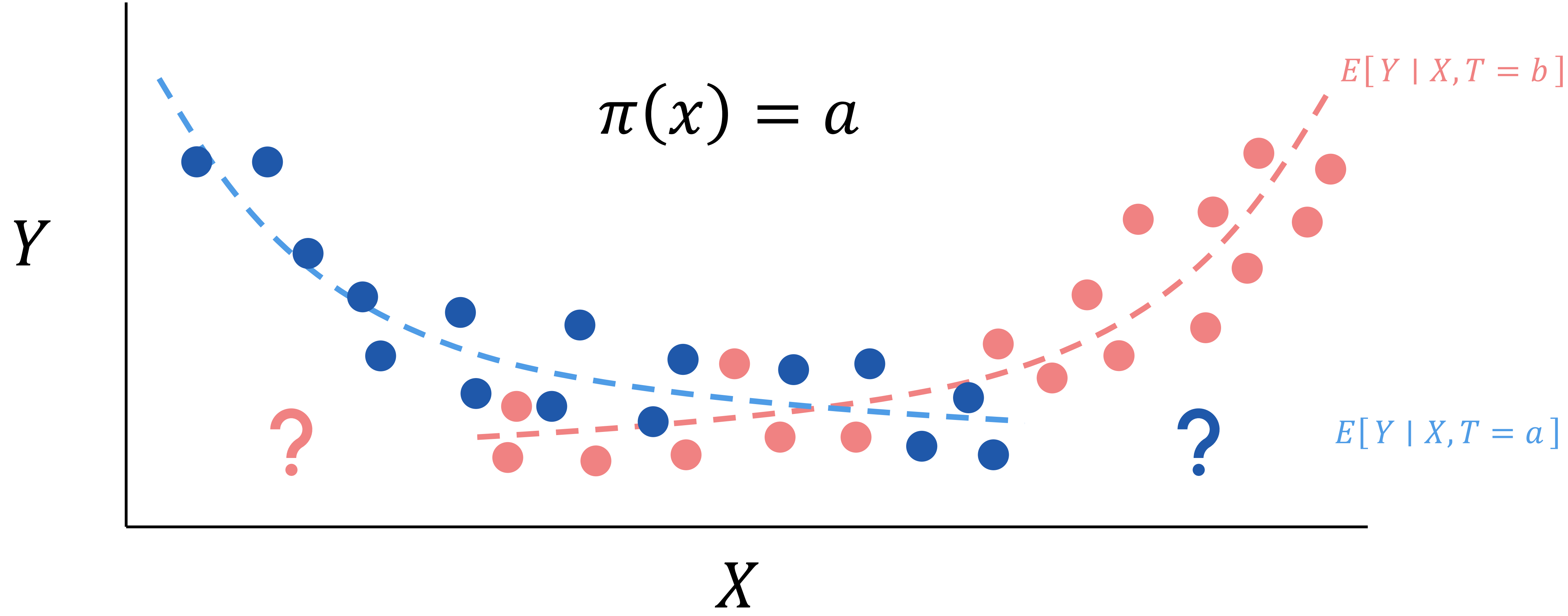
Which policies can we evaluate?

Coverage: For all x, $\pi(x) = t$ ⟶ $p(t \mid x) > 0$



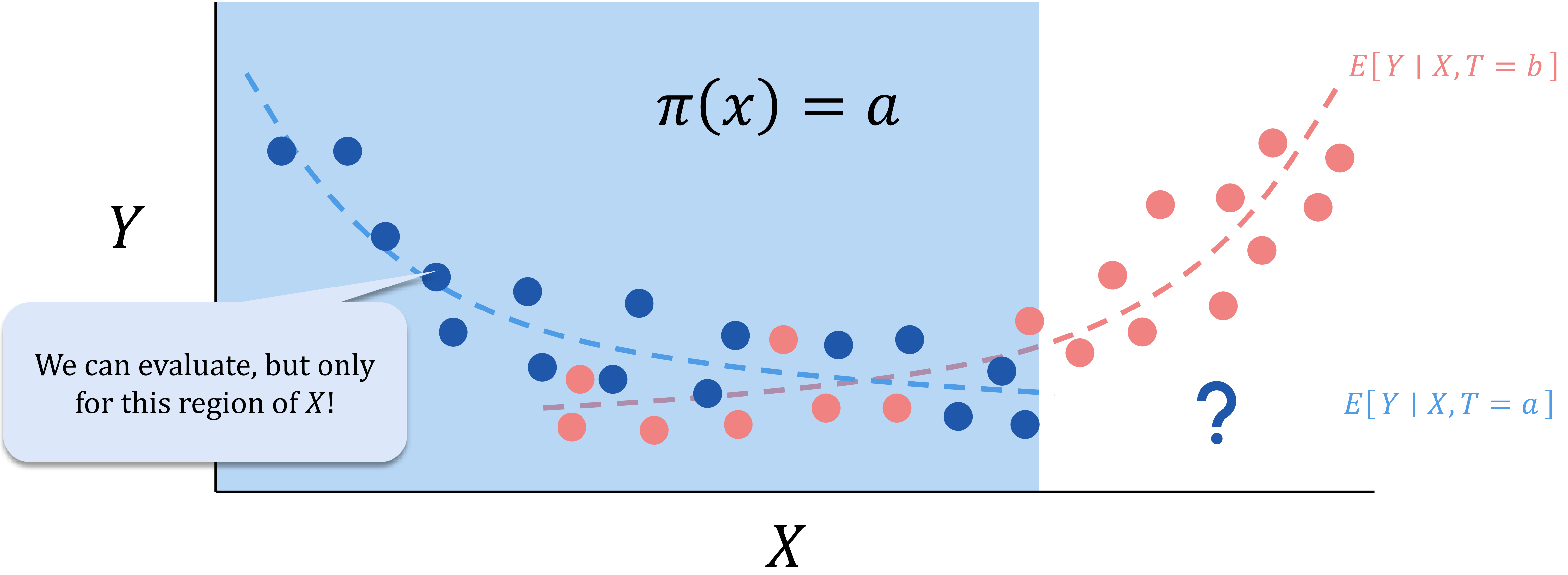
Where can we evaluate a fixed policy?

Coverage: For all x, $\pi(x) = t$ ⟶ $p(t \mid x) > 0$



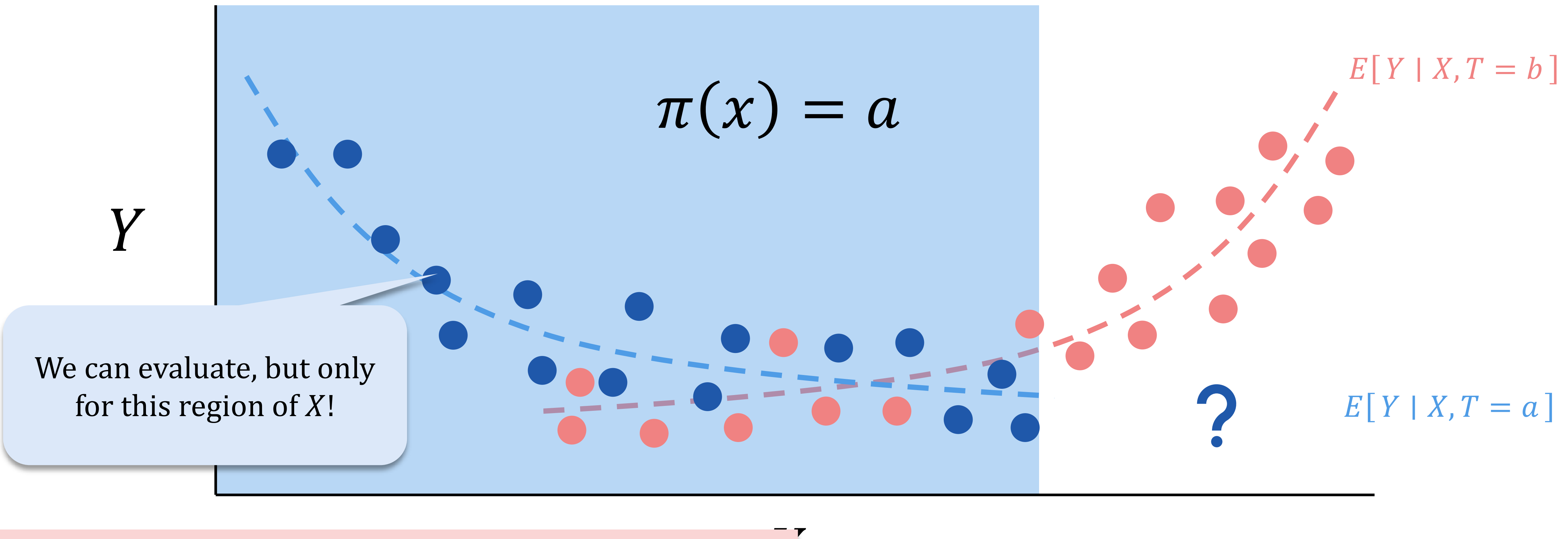
Where can we evaluate a fixed policy?

Coverage: For all x, $\pi(x) = t$ ⟶ $p(t \mid x) > 0$



Where can we evaluate a fixed policy?

Coverage: For all x, $\pi(x) = t$ ⟶ $p(t \mid x) > 0$



Characterization of Overlap in Observational Studies
AISTATS 2020
Oberst, M.*, Johansson, F.*, Wei, D.*, Gao, T., Brat, G., Sontag, D., & Varshney, K.

Goals

- By the end of this primer, you should understand…
  - What a **causal effect** is, and when we can infer it from observational data

    "What is the effect of treatment vs. placebo?"
  - How **causal effect estimation** is a special case of **policy evaluation**

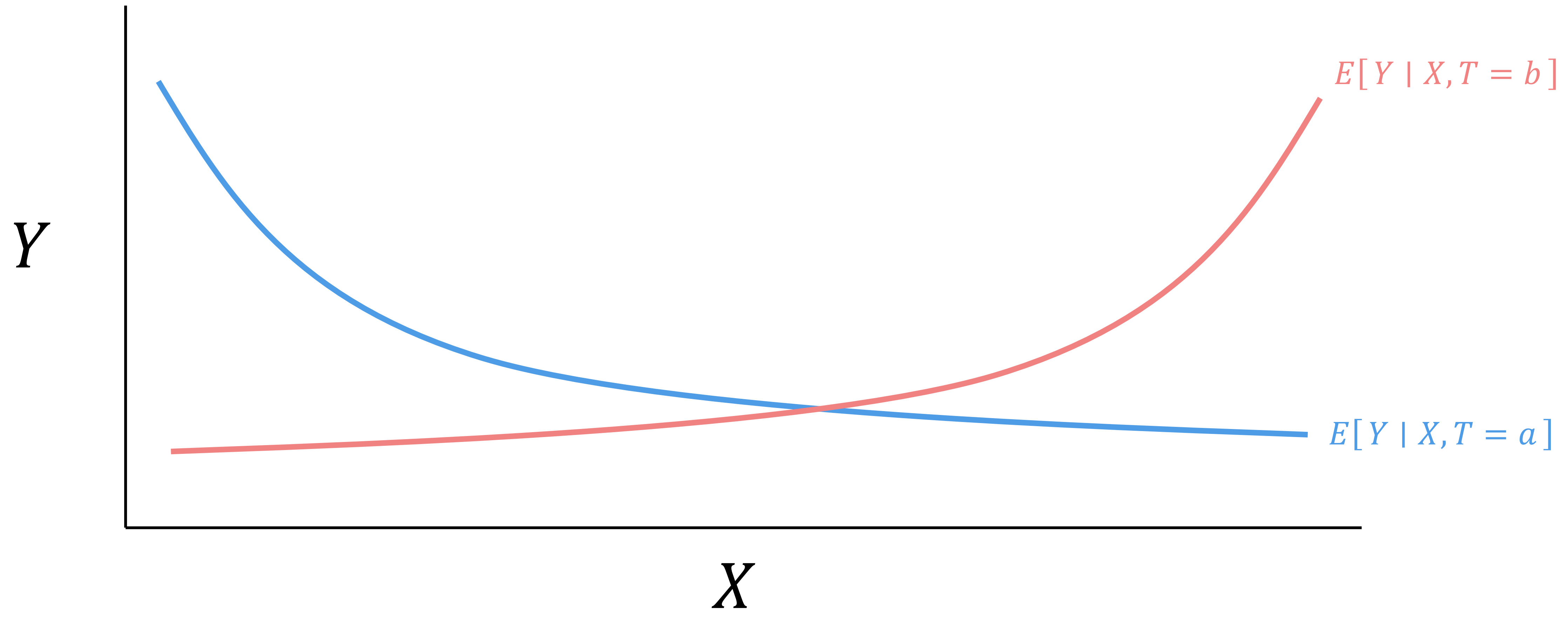
    "What would happen if I implemented a new set of guidelines?"
  - How to use policy evaluation to **learn optimal policies**

    "What is the optimal treatment strategy?"

Learning Policies from Data

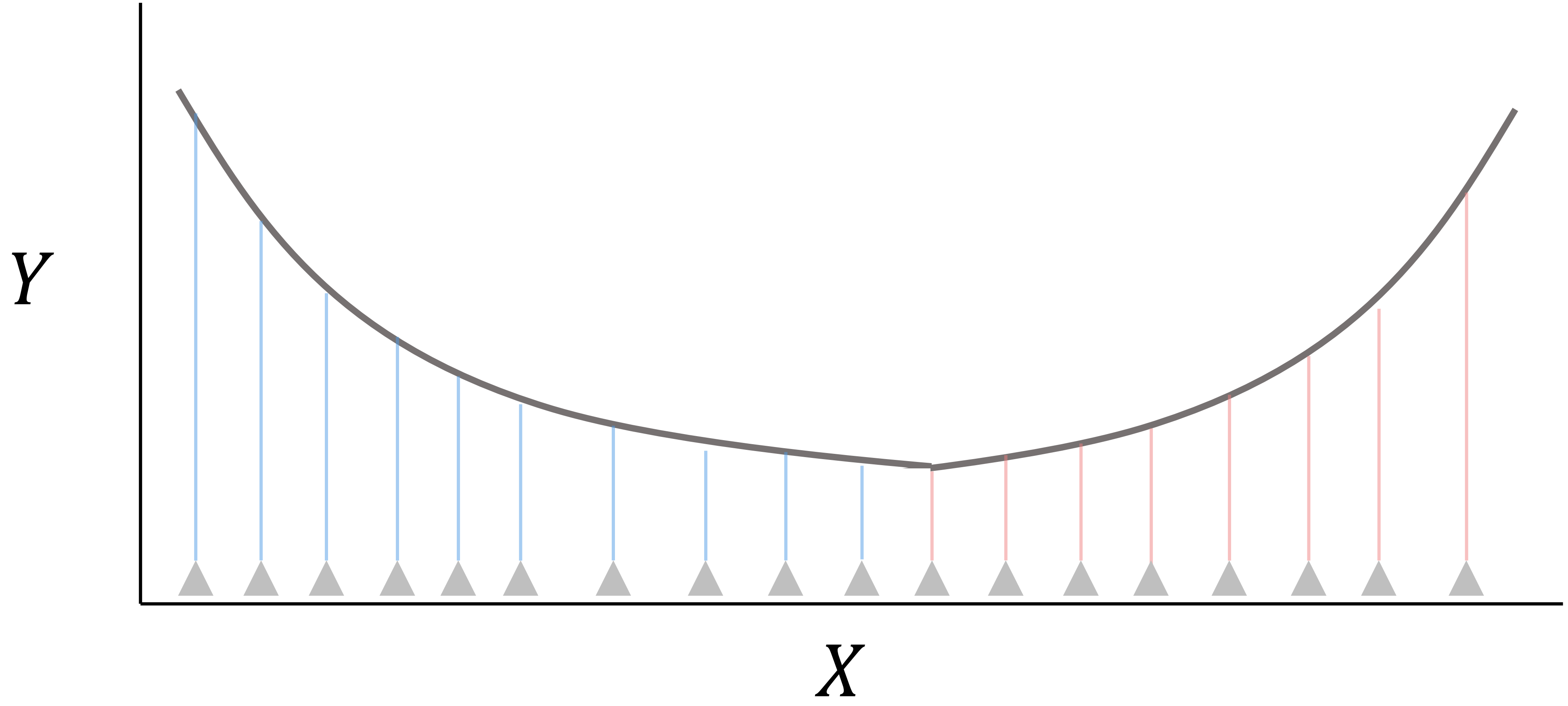
What is the optimal policy?

$$\mathrm{E}\left[Y\left(\pi(x)\right)\right] = E_x\left[E[\,Y \mid X, T = \pi(x)\,]\right]$$



What is the optimal policy?

$$\mathrm{E}\big[Y\big(\pi(x)\big)\big] = E_x\big[E[\,Y \mid X, T = \pi(x)\,]\big]$$



How to find the optimal policy?

Option #1: Estimate outcomes for each action, choose the best one!

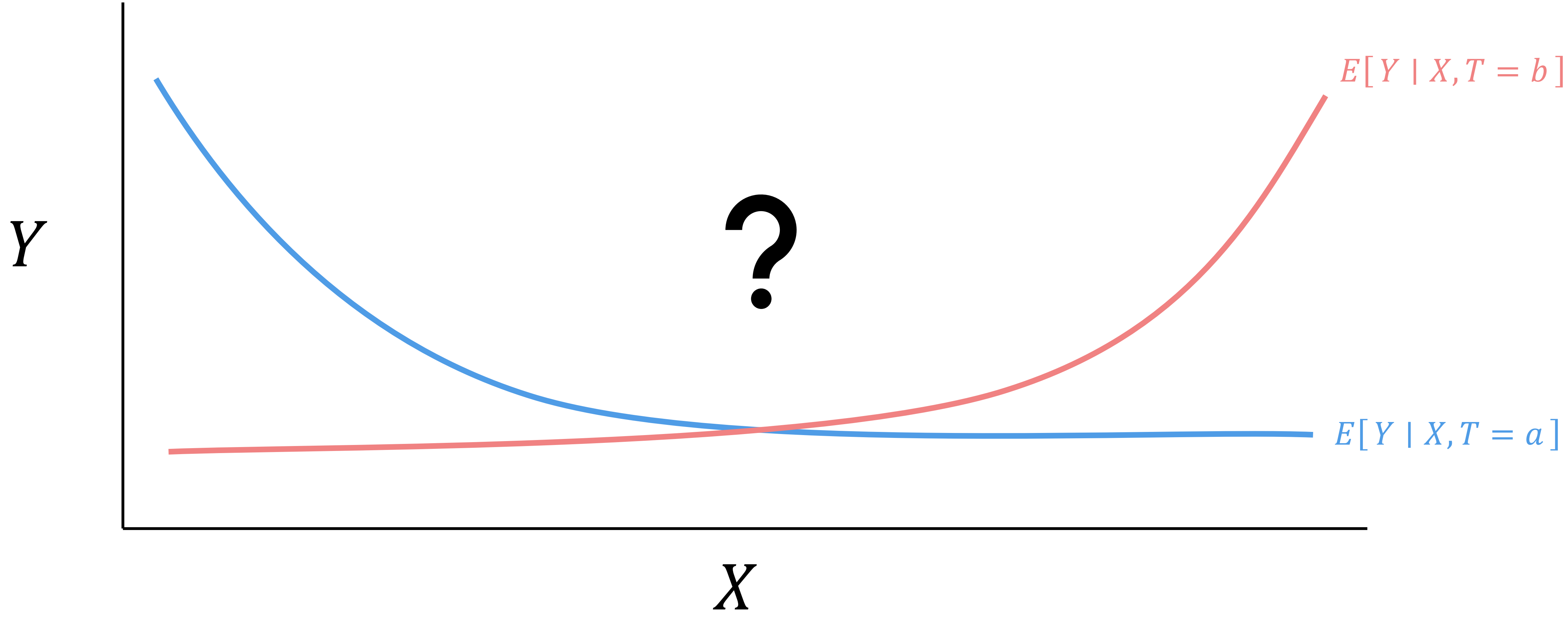
$$\pi(x) := \text{argmax}_{t \in T}\, f(x, t)$$

- “Indirect” method
- Requires that regression function be “correct”
- The resulting policy could also be quite complex

$f(x, t) \approx E[Y \mid X = x, T = t]$

Indirect method may yield complex policies

$\pi(x) = \text{argmax}_{t \in T} \, f(x,t)$

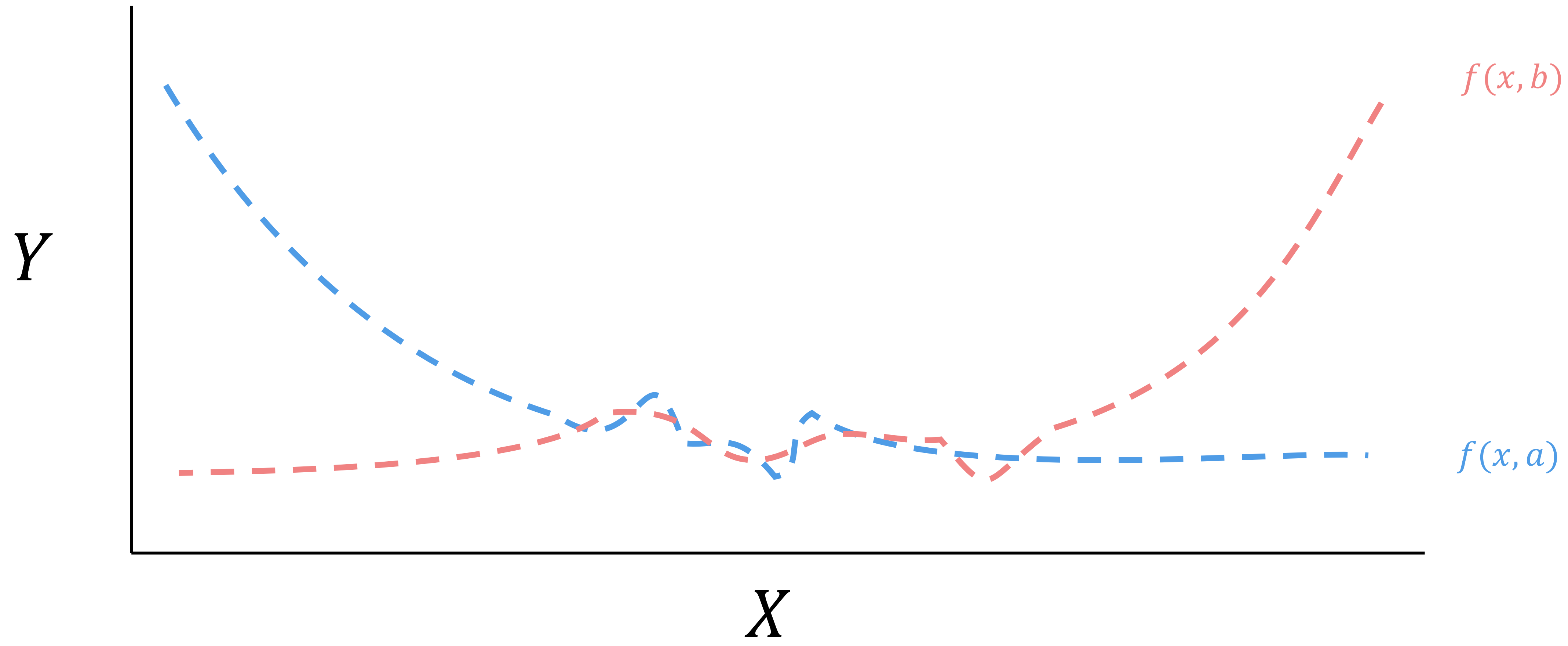


Indirect method may yield complex policies

$$\overline{\pi(x)} = \text{argmax}_{t \in T} f(x, t)$$

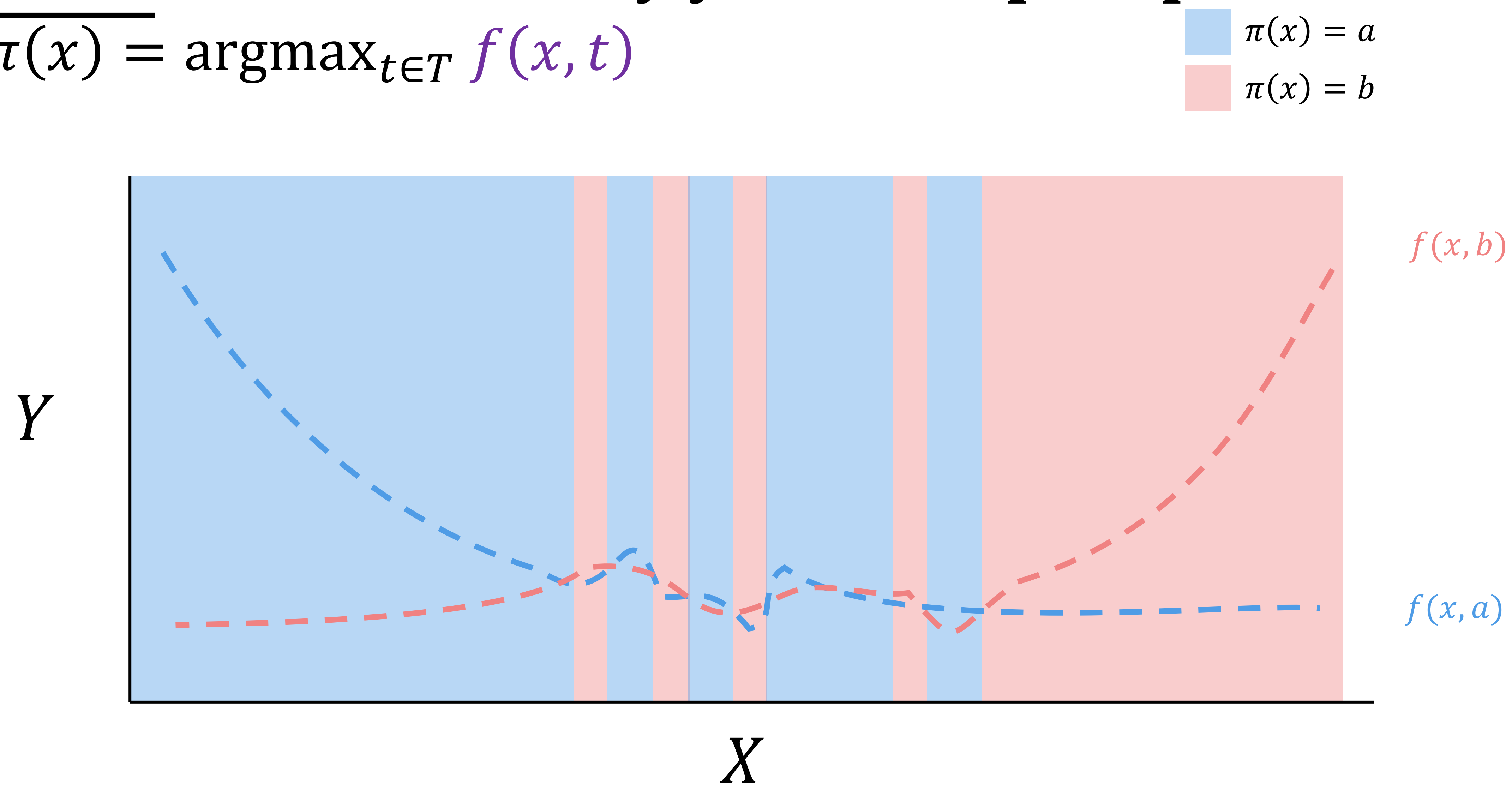
$\pi(x) = a$

$\pi(x) = b$



Indirect method may yield complex policies

$$\overline{\pi(x)} = \text{argmax}_{t \in T}\, f(x,t)$$



How to find the optimal policy?

Option #2: Directly solve for the best policy in a certain class

$$\pi^* = \text{argmax}_{\pi \in \Pi} \hat{E}\left[Y\big(\pi(x)\big)\right]$$

Now we are **restricting** to some **policy class** Π (e.g., linear treatment rules)

How to find the optimal policy?

Option #2 ("Direct"): Solve for the best policy in a certain class

$$\pi^* = \text{argmax}_{\pi \in \Pi} \frac{1}{n} \sum_{i=1}^{n} \sum_{a \in A} 1\{\pi(x_i) = a\} * \widehat{W}_a(i)$$

Now we are **restricting** to some **policy class** Π (e.g., linear treatment rules)

"Doubly Robust" Estimator

$$\widehat{W}_a(i) := \frac{1\{t_i = a\}}{e(x_i, t_i)} (Y_i - f(x_i, a) + f(x_i, a)$$

Directly optimizing a policy

Option #2 (“Direct”): Solve for the best policy in a certain class

$$\pi^* = \text{argmax}_{\pi \in \Pi} \frac{1}{n} \sum_{i=1}^{n} \sum_{a \in A} 1\{\pi(x_i) = a\} * \widehat{W}_a(i)$$

- We can optimize over simple (interpretable) policies, even if outcome / propensity models are very flexible and complex
- Conceptually straightforward to add other actions (e.g., deferral)

Not obvious how to optimize this! Why?

Policy Optimization as classification

$$\pi^* = \text{argmax}_{\pi \in \Pi} \frac{1}{n} \sum_{i=1}^{n} \sum_{a \in A} 1\{\pi(x_i) = a\} * \widehat{W}_a(i)$$

Difficult optimization problem! Analogous to weighted 0-1 loss in classification.

Policy Optimization as classification

Use convex surrogate for indicator function

$$\pi^* = \text{argmax}_{\pi \in \Pi} \frac{1}{n} \sum_{i=1}^{n} \sum_{a \in A} 1\{\pi(x_i) = a\} * \widehat{W}_a(i)$$

Convert to minimization

Redefine weights (e.g., ensure nonnegative) so that reformulated problem remains convex

Goal: Reformulate optimization problem so that

- It is easier to optimize
- It has the same optimal solution

Goals

- By the end of this primer, you should understand…
  - What a **causal effect** is, and when we can infer it from observational data

    "What is the effect of treatment vs. placebo?"
  - How **causal effect estimation** is a special case of **policy evaluation**

    "What would happen if I implemented a new set of guidelines?"
  - How to use policy evaluation to **learn optimal policies**

    "What is the optimal treatment strategy?"

Recommended Reading (Policy Learning)

[1] Y. Q. Zhao, E. B. Laber, Y. Ning, S. Saha, and B. E. Sands, **Efficient augmentation and relaxation learning for individualized treatment rules using observational data,** J. Mach. Learn. Res., vol. 20, pp. 1–23, 2019.

[2] X. Huang, Y. Goldberg, and J. Xu, **Multicategory individualized treatment regime using outcome weighted learning,** Biometrics, August 2018, pp. 1216–1227, 2019.

[3] S. Athey and S. Wager, **Policy Learning with Observational Data**, Econometrica (Forthcoming), 2020.

[4] N. Kallus, **More Efficient Policy Learning via Optimal Retargeting**, J. Am. Stat. Assoc., 2020.

Disclaimer: Selected set of (very) recent papers, but good place to find references to broader literature