Better Lee Bounds

Vira Semenova

July, 29, 2021

Introduction

Endogenous selection $\Rightarrow$ need bounds on ATE.

Introduction

Endogenous selection $\Rightarrow$ need bounds on ATE.

Lee (2009): covariates X map to a class of bounds on ATE

- any covariate subset $X' \subseteq X$ maps to a pair of valid bounds
- the full set X maps to sharp (the tightest possible) bounds

Introduction

Endogenous selection $\Rightarrow$ need bounds on ATE.

Lee (2009): covariates X map to a class of bounds on ATE

- any covariate subset $X' \subseteq X$ maps to a pair of valid bounds
- the full set X maps to sharp (the tightest possible) bounds

Sharp bounds are difficult to estimate, non-sharp bounds may not be very useful.

Outline

1. Overview of Lee (2009) bounds
2. Better Lee Bounds
3. JobCorps revisited

Lee (2009): notation

Potential employment and wage outcomes

- $D = 1$ if subject wins a lottery
- $S(d) = 1$ if employed when $D = d$ for $d \in \{1, 0\}$
- $Y(d)$ wage when $D = d$ for $d \in \{1, 0\}$

Lee (2009): notation

Potential employment and wage outcomes

- $D = 1$ if subject wins a lottery
- $S(d) = 1$ if employed when $D = d$ for $d \in \{1, 0\}$
- $Y(d)$ wage when $D = d$ for $d \in \{1, 0\}$

Observed data are $(X, D, S, S \cdot Y)$. Y exists only if $S = 1$.

Lee (2009): notation

Potential employment and wage outcomes

- $D = 1$ if subject wins a lottery
- $S(d) = 1$ if employed when $D = d$ for $d \in \{1, 0\}$
- $Y(d)$ wage when $D = d$ for $d \in \{1, 0\}$

Observed data are $(X, D, S, S \cdot Y)$. Y exists only if $S = 1$.

Target population

$$S(1) = S(0) = 1 \quad \text{a.s.}$$

Lee (2009): notation

Potential employment and wage outcomes

- $D = 1$ if subject wins a lottery
- $S(d) = 1$ if employed when $D = d$ for $d \in \{1, 0\}$
- $Y(d)$ wage when $D = d$ for $d \in \{1, 0\}$

Observed data are $(X, D, S, S \cdot Y)$. Y exists only if $S = 1$.

Target population

$$S(1) = S(0) = 1 \quad \text{a.s.}$$

Target parameter

$$\text{ATE} = \mathbb{E}[Y(1) - Y(0) \mid S(1) = S(0) = 1]$$

Lee's monotonicity assumption

$$S(0) = 1 \Rightarrow S(1) = 1.$$

Lee (2009): basic bounds

$\mathbb{E}[Y(0) \mid S(0) = 1]$ is point-identified. $\mathbb{E}[Y(1) \mid S(0) = 1]$ is not.

Lee (2009): basic bounds

$\mathbb{E}[Y(0) \mid S(0) = 1]$ is point-identified. $\mathbb{E}[Y(1) \mid S(0) = 1]$ is not.

Always-takers' share among the treated

$$p_0 = \frac{\Pr[S = 1|D = 0]}{\Pr[S = 1|D = 1]} \in (0, 1)$$

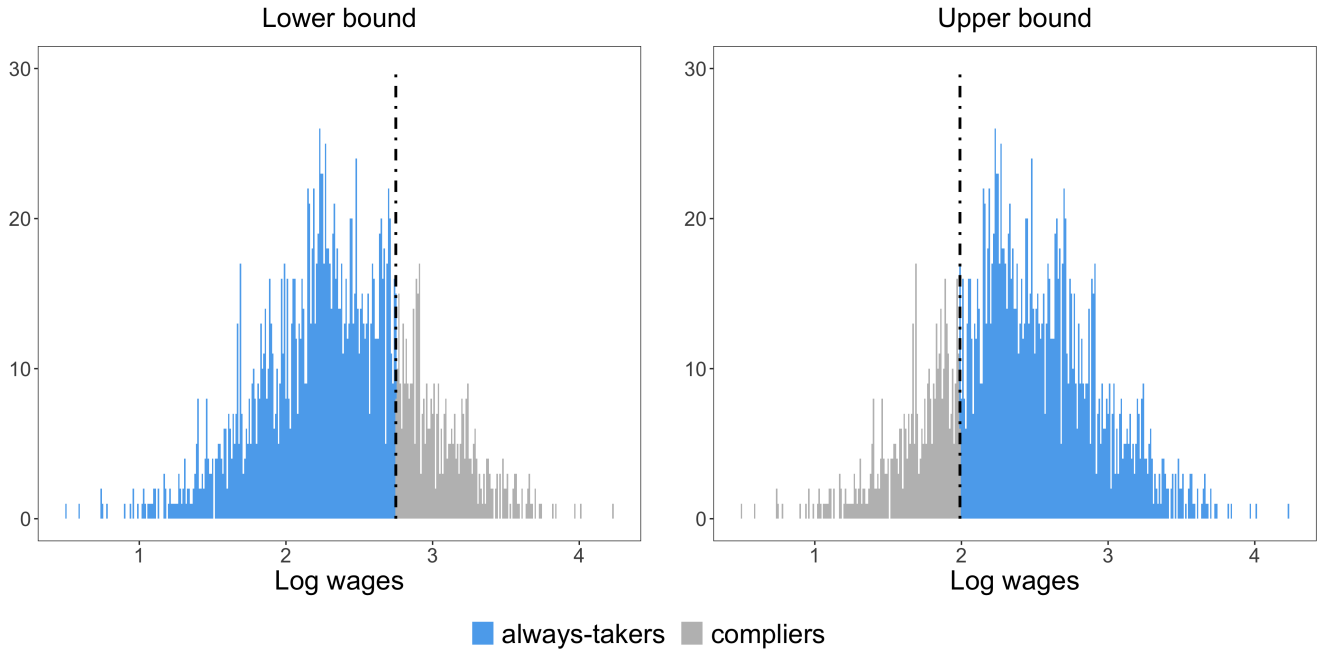
Lee (2009): basic bounds

$\mathbb{E}[Y(0) \mid S(0) = 1]$ is point-identified. $\mathbb{E}[Y(1) \mid S(0) = 1]$ is not.

Always-takers' share among the treated

$$p_0 = \frac{\Pr[S = 1|D = 0]}{\Pr[S = 1|D = 1]} \in (0, 1)$$

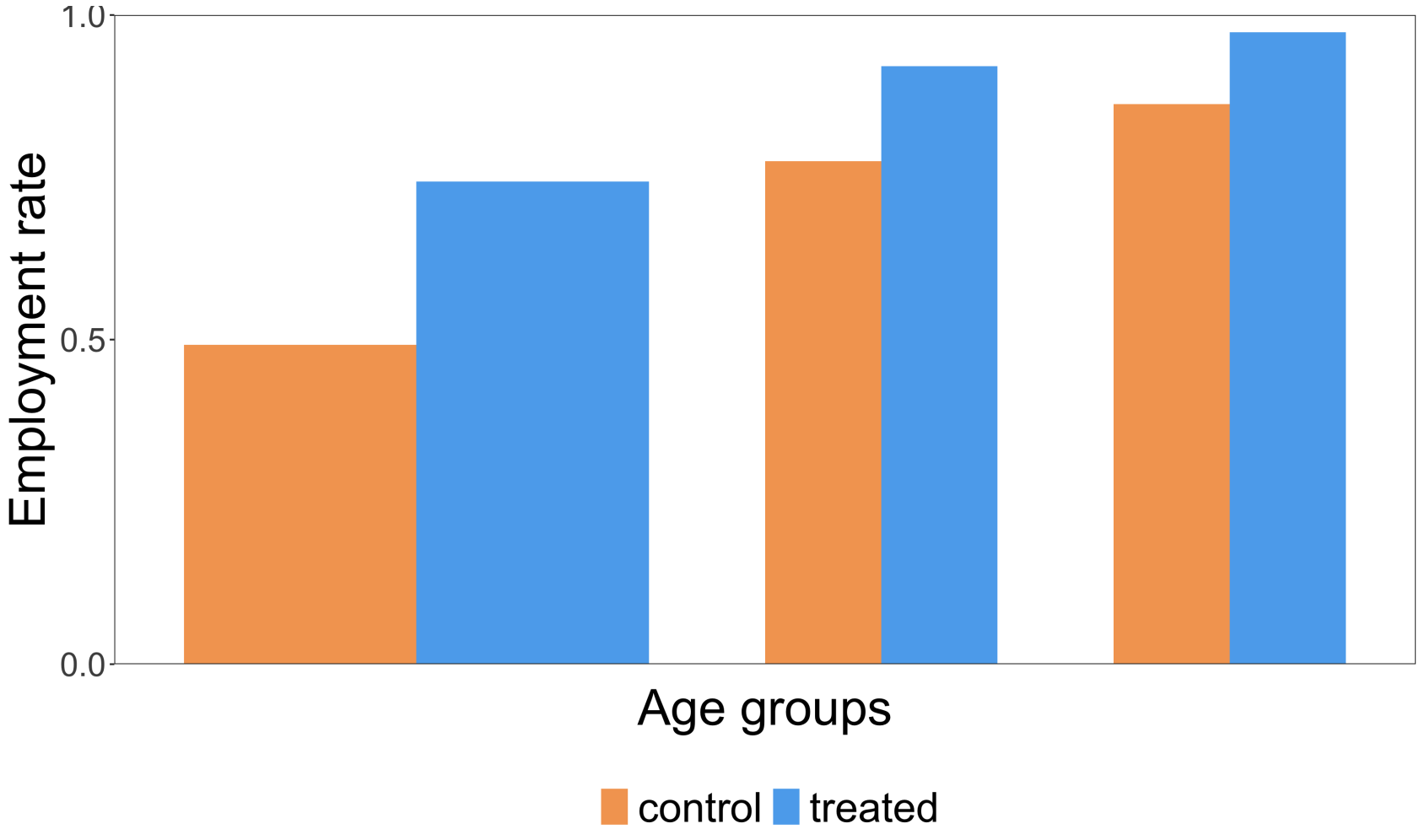
Bounds on $\mathbb{E}[Y(1) \mid S(0) = 1]$ are trimmed means:



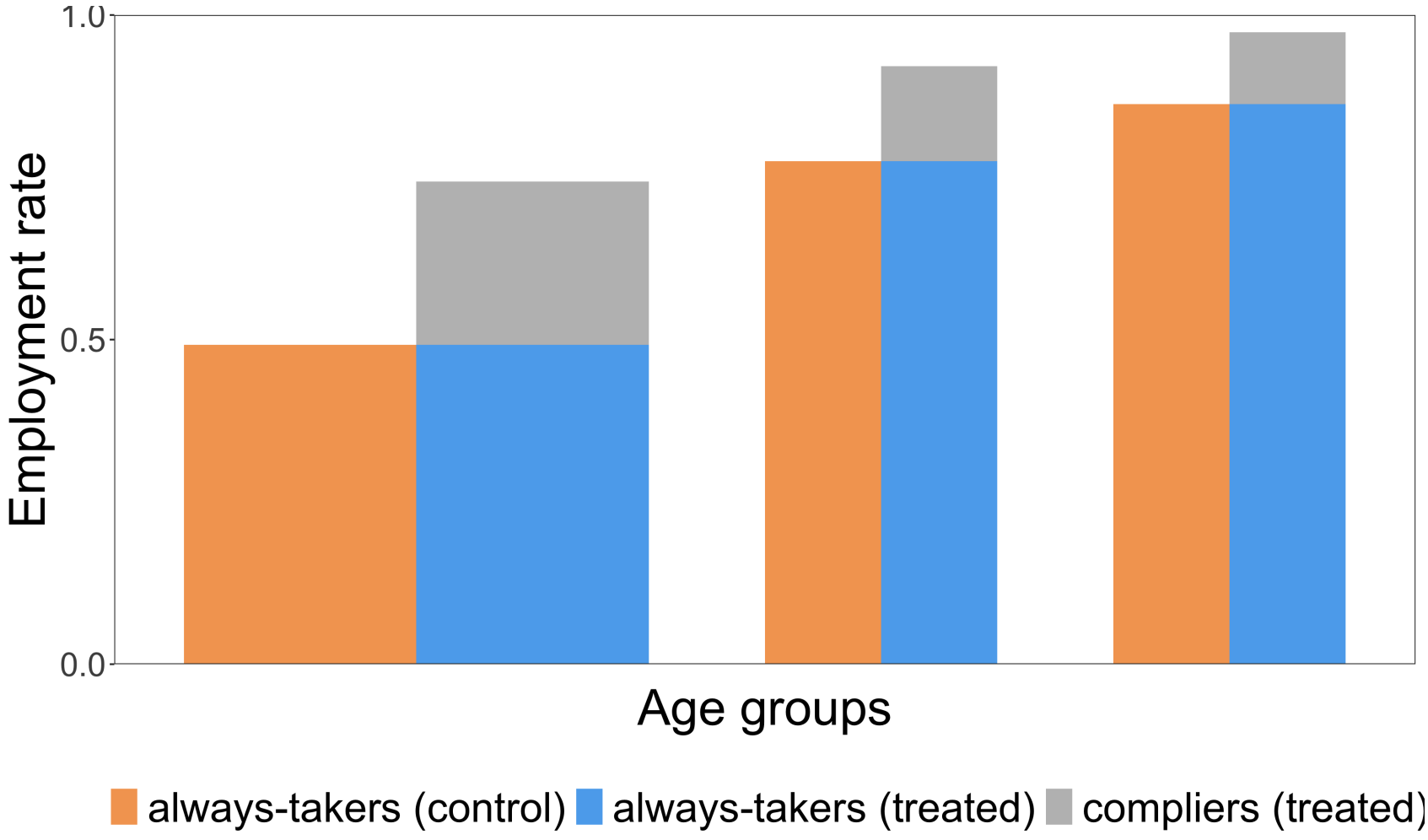
$p_0 = 0.77$, quant$(p_0) = 2.75$

lower bound $\widehat{\beta}_{\text{basic}} = 2.178 - 1.38 = 0.78$

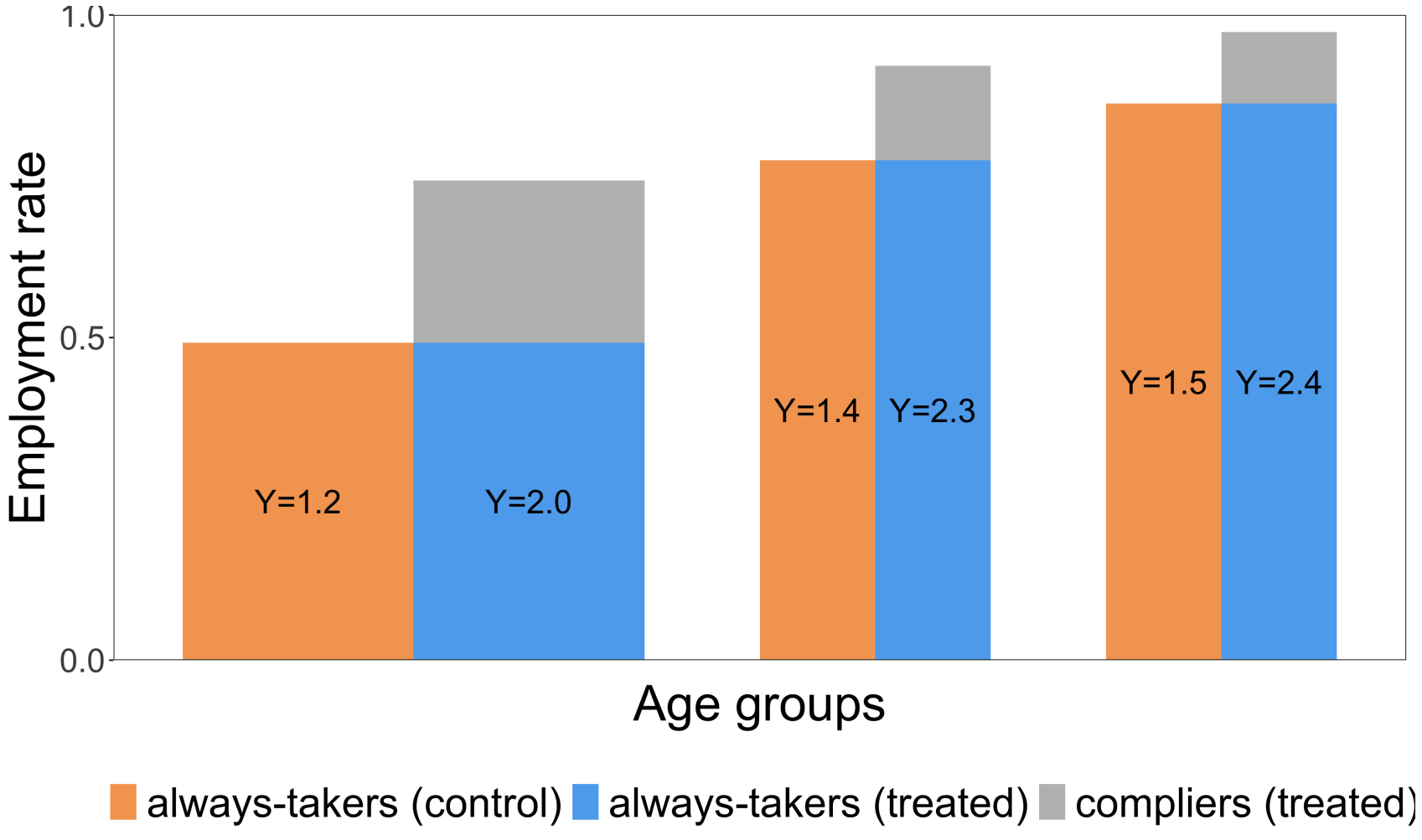
Lee (2009): covariate-based bounds



Lee (2009): covariate-based bounds, cont.

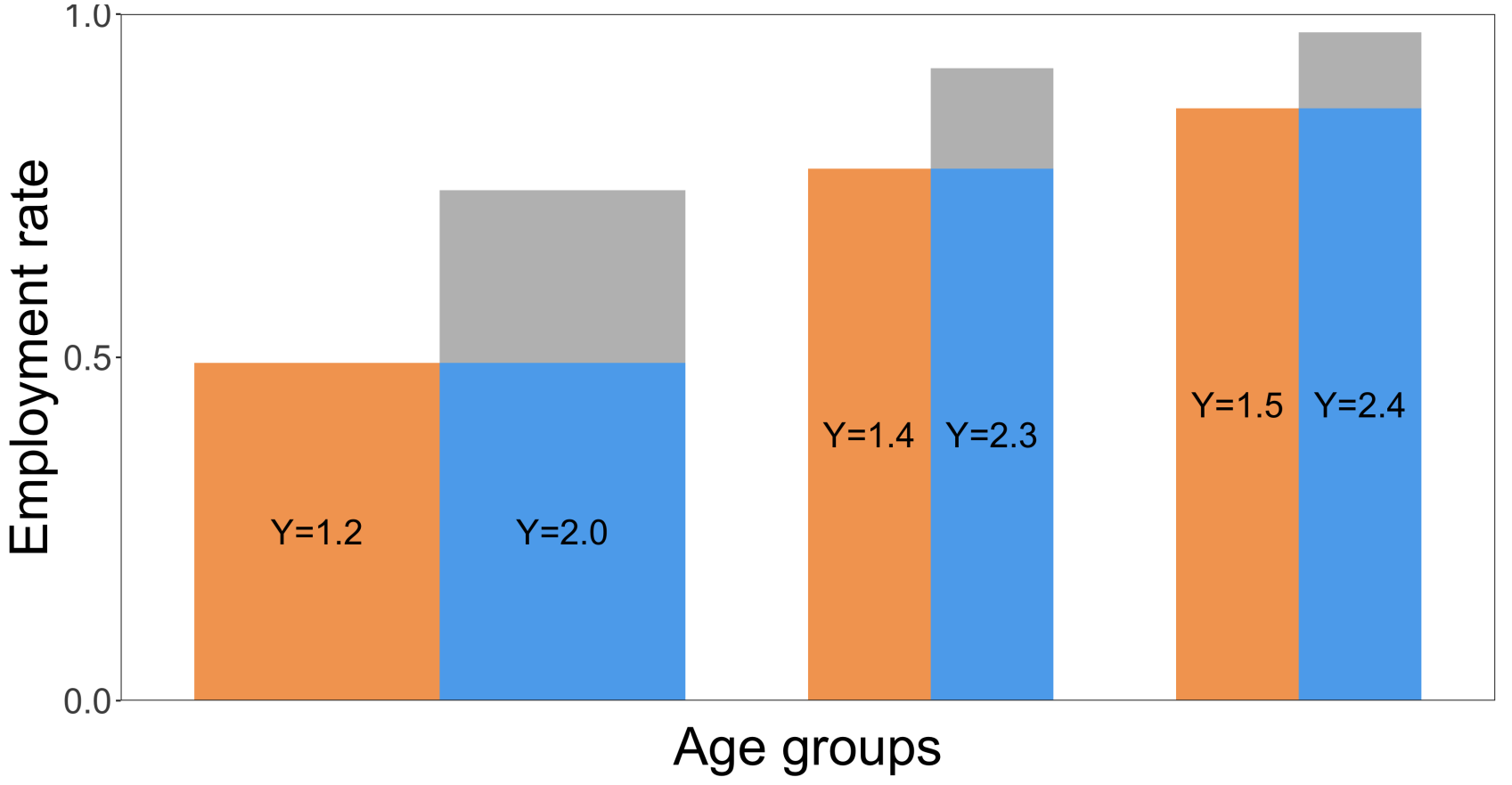


Lee (2009): covariate-based bounds, cont.



lower bound $\widehat{\beta}_{\text{discrete}} = 1/2(2.0 - 1.2) + 1/4(2.3 - 1.4) + 1/4(2.4 - 1.5) = 0.85$

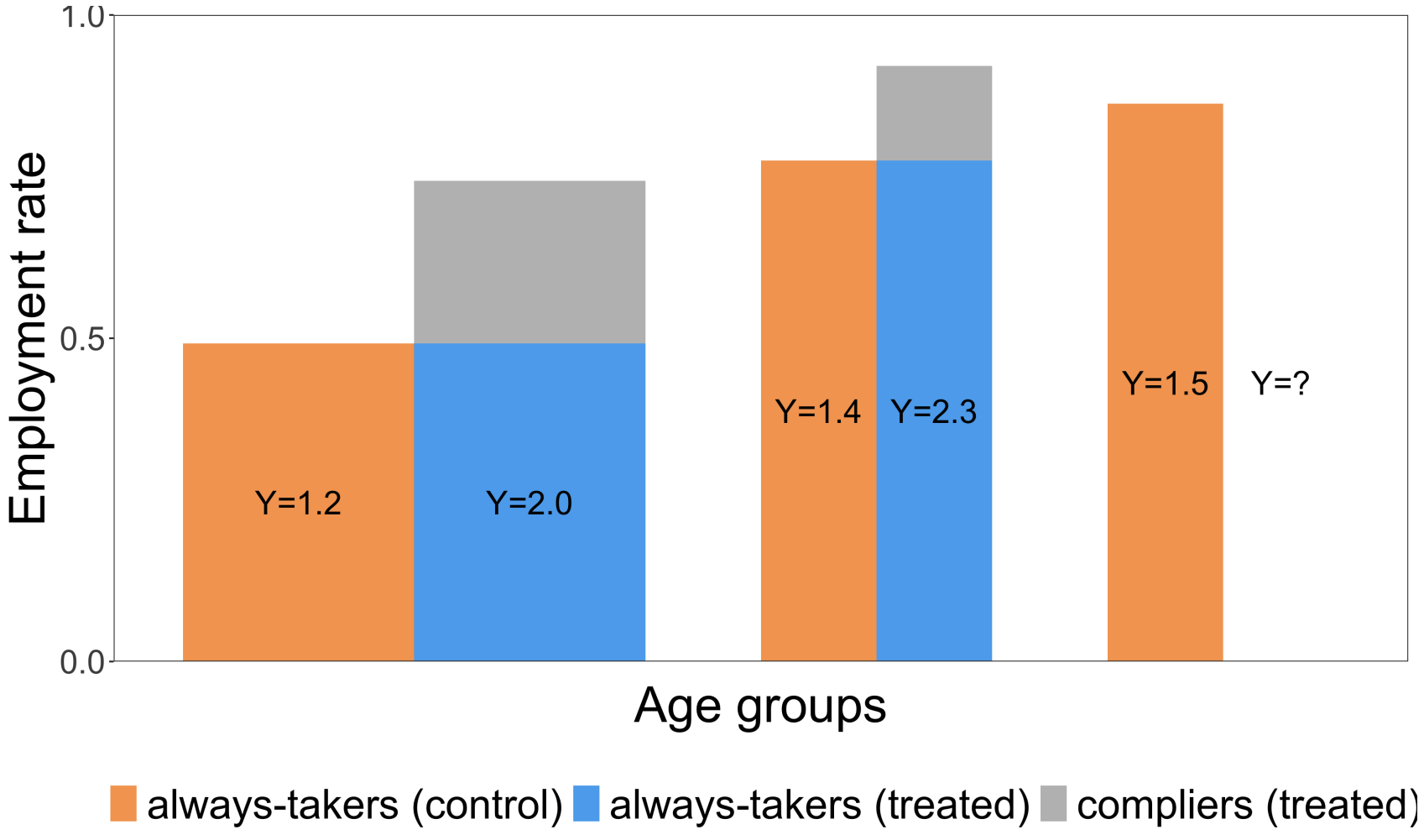
Lee (2009): covariate-based bound, cont.



$$\widehat{\beta}_{\text{discrete}} = 1/2(2.0 - 1.2) + 1/4(2.3 - 1.4) + 1/4(2.4 - 1.5) = 0.85$$

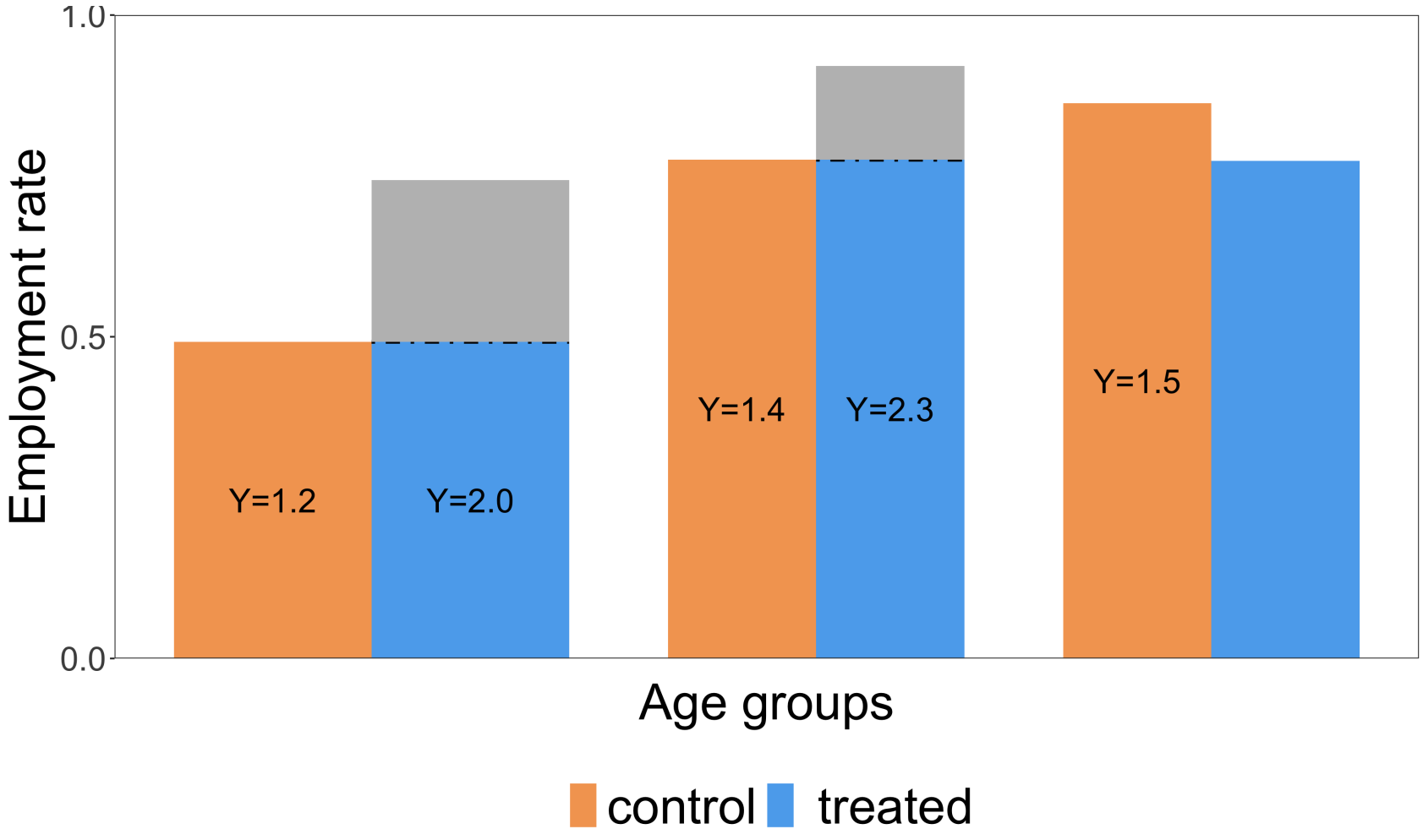
$$\widehat{\beta}_{\text{discrete}} = 0.85 > \widehat{\beta}_{\text{basic}} = 0.78$$

Lee (2009): degenerate cell example



$$\widehat{\beta}_{\text{discrete}} = {}^{1}\!/_{2}(2.0 - 1.2) + {}^{1}\!/_{4}(2.3 - 1.4) + {}^{1}\!/_{4}(? - 1.5) = ?$$

Lee (2009): heterogeneous monotonicity exmaple



$$\widehat{\beta}_{\text{discrete}} = 1/2(2.0 - 1.2) + 1/4(2.3 - 1.4) + ? = ?$$

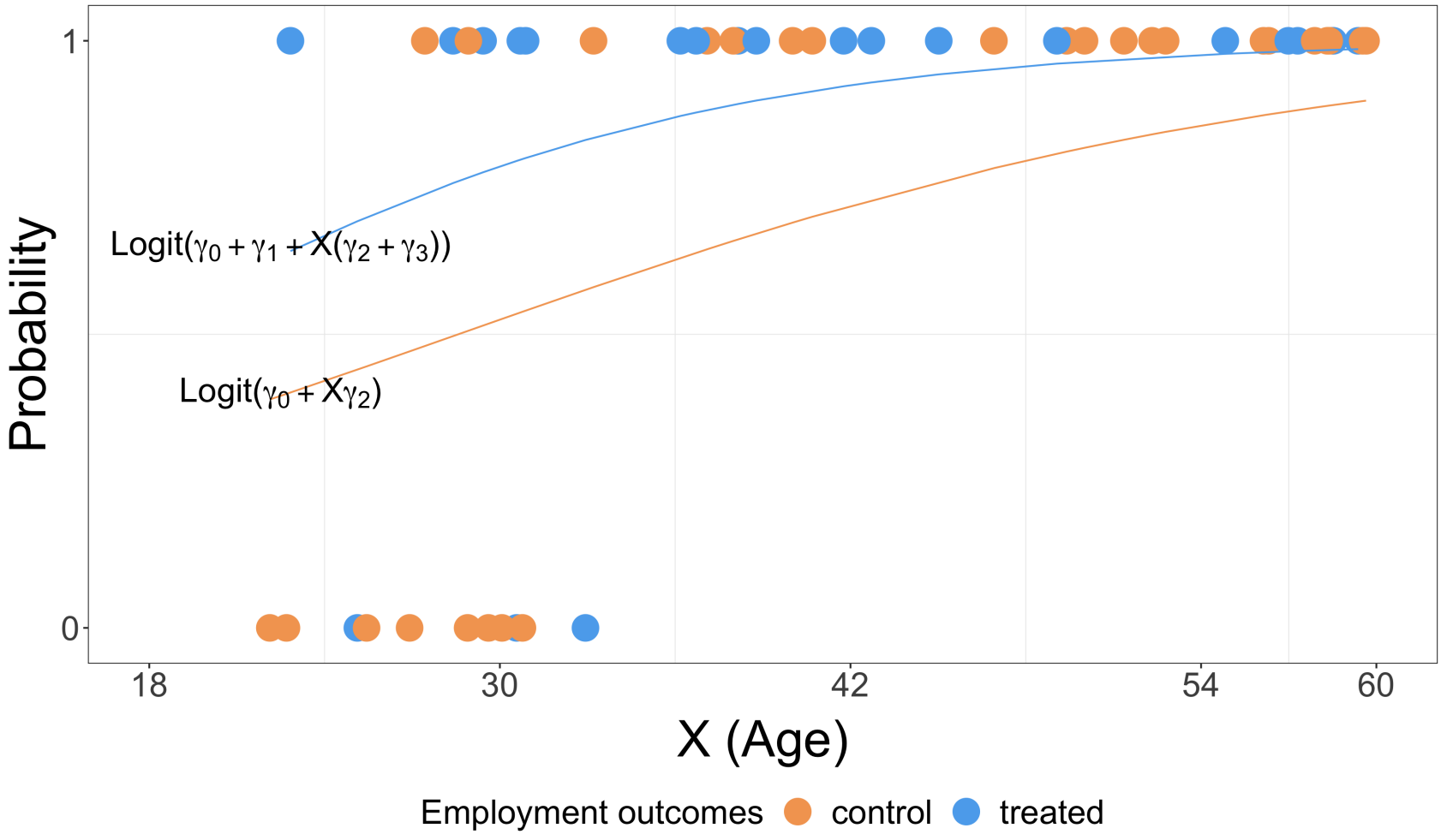
Outline

1. Overview of Lee (2009) bounds
2. Better Lee Bounds
3. JobCorps revisited

Better Lee Bounds: the probability of employment

employment probability

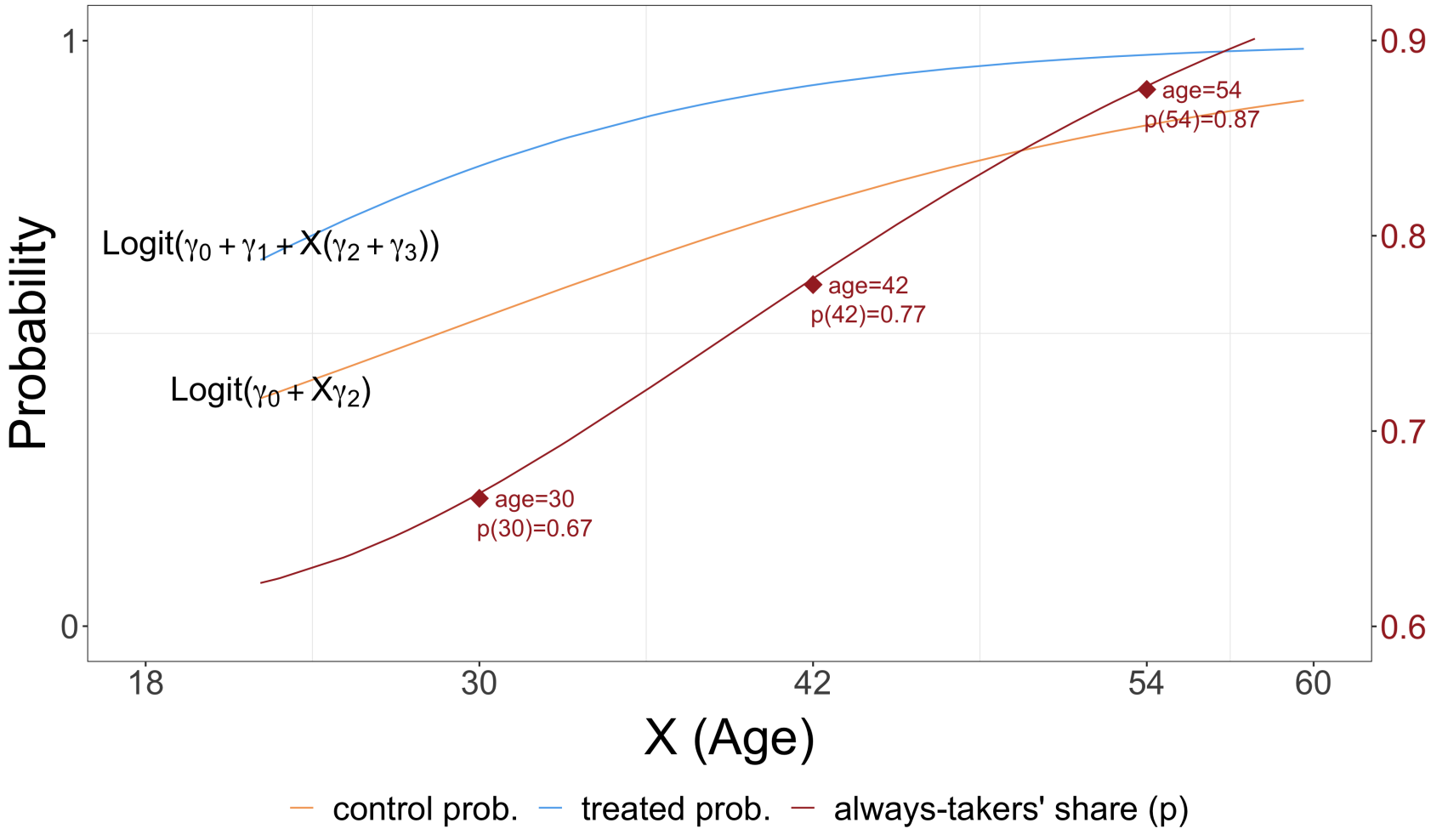
$$\Pr(S = 1 \mid D, X) = \text{Logit}(\gamma_0 + D \cdot \gamma_1 + X \cdot \gamma_2 + D \cdot X \cdot \gamma_3)$$



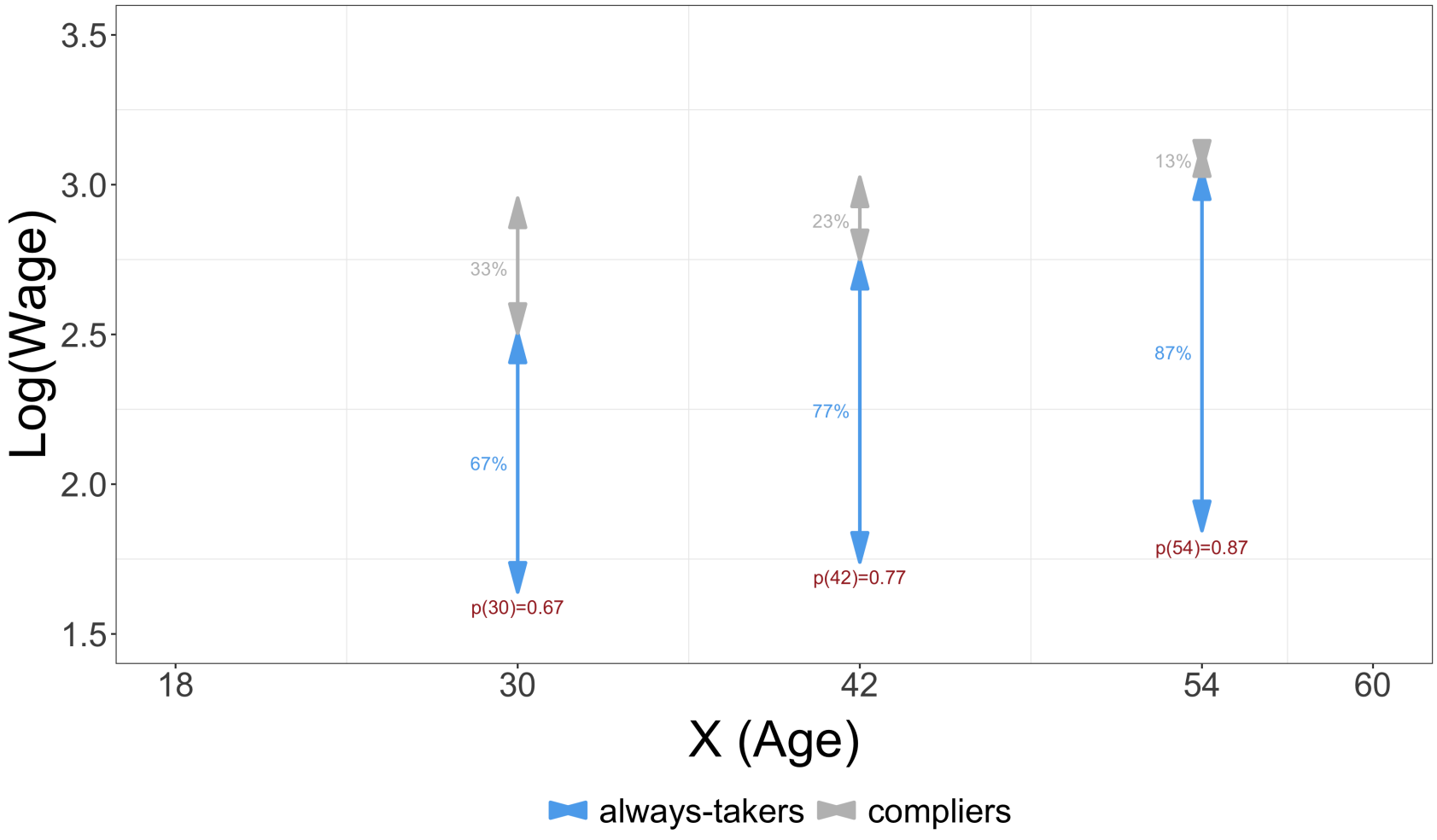
Better Lee Bounds: the always-takers' share among the treated

always-takers' share among the treated

$$p(X) = \frac{\text{Logit}(\gamma_0 + X \cdot \gamma_2)}{\text{Logit}(\gamma_0 + \gamma_1 + X \cdot (\gamma_2 + \gamma_3))}$$



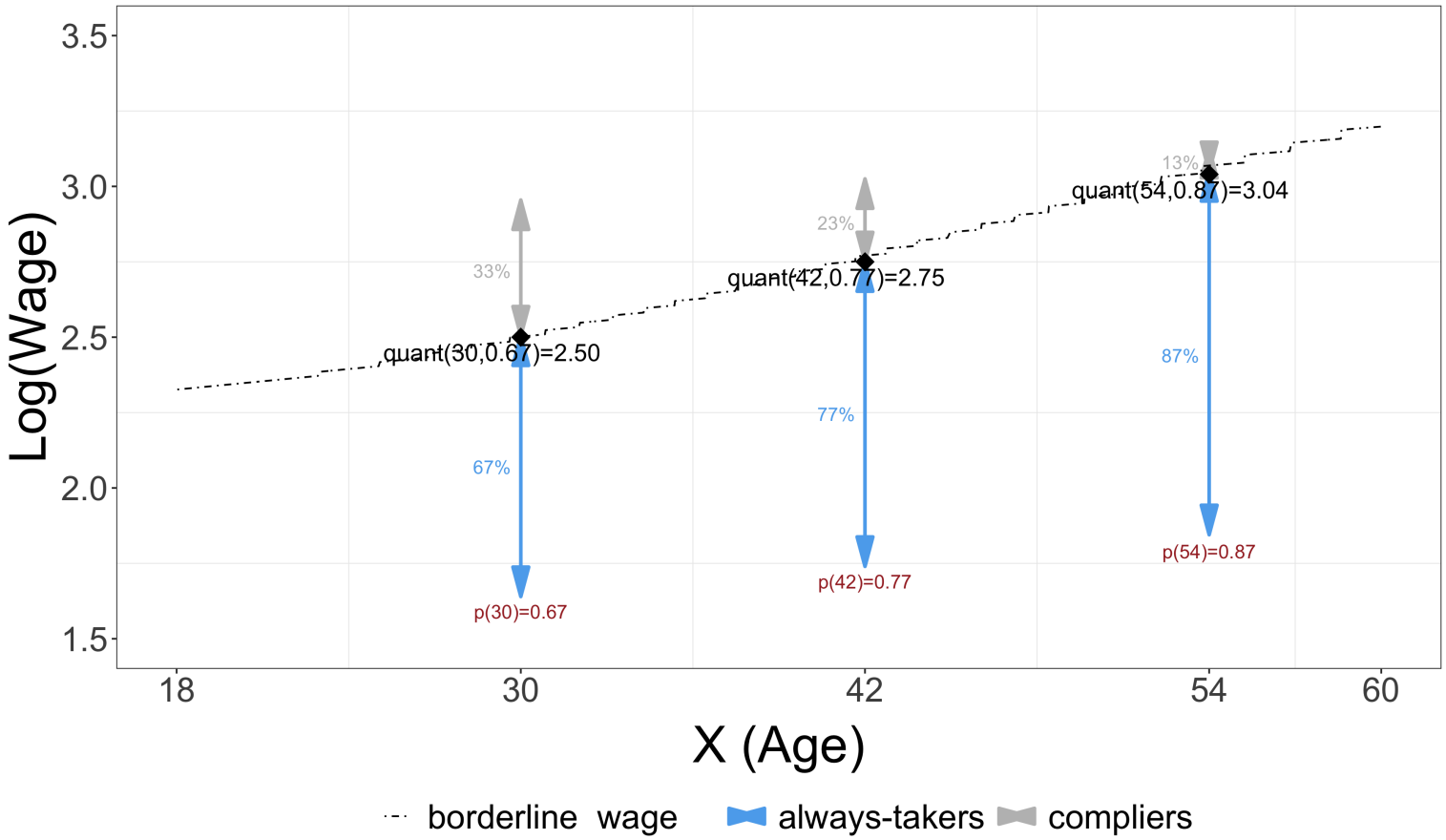
Better Lee Bounds: the worst case for the always-takers



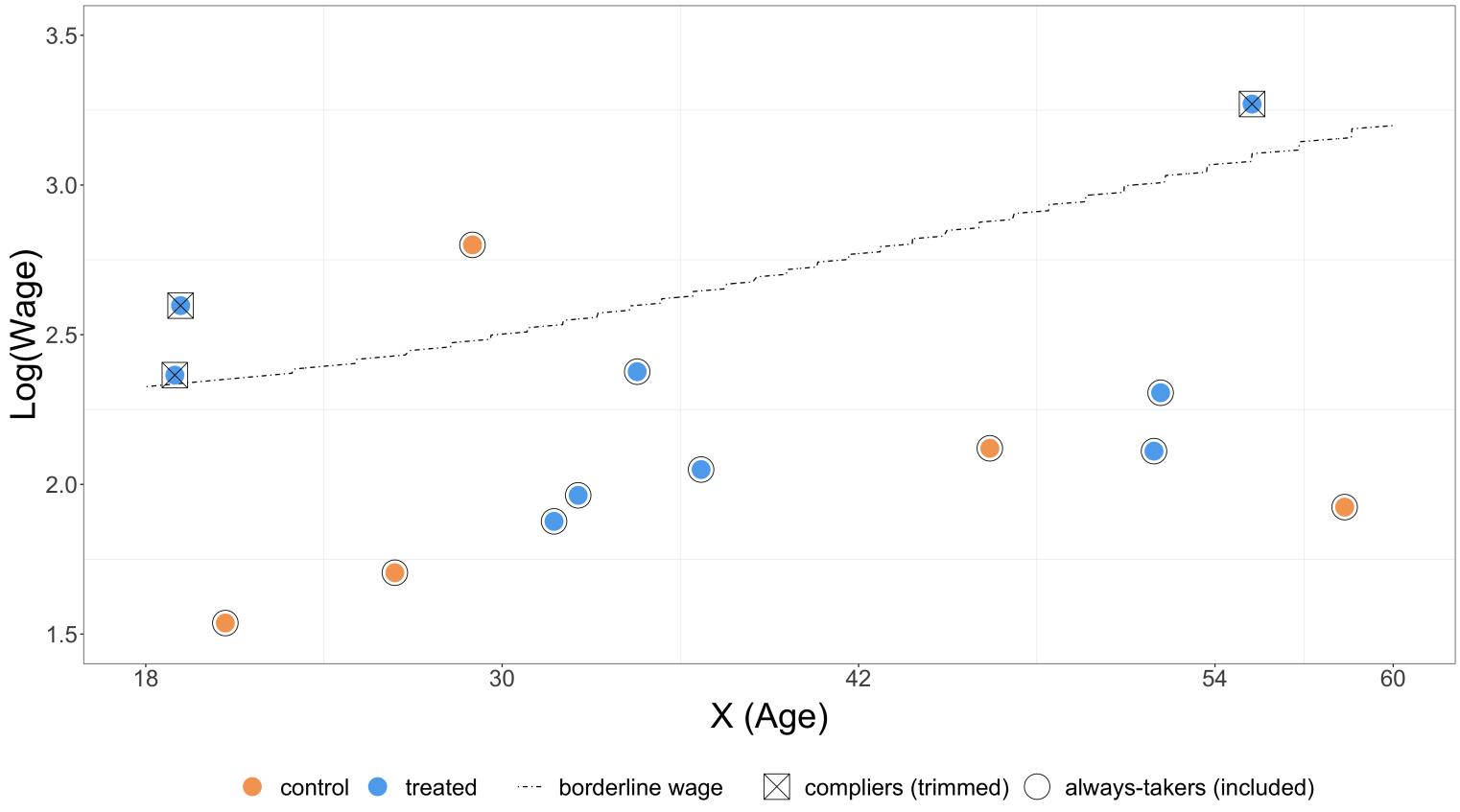
Better Lee Bounds: the borderline always-takers' wage

quant(X, u) is u-quantile of the treated wages

the borderline wage is quant($X, p(X)$)



Better Lee Bounds: the trimmed wage sample



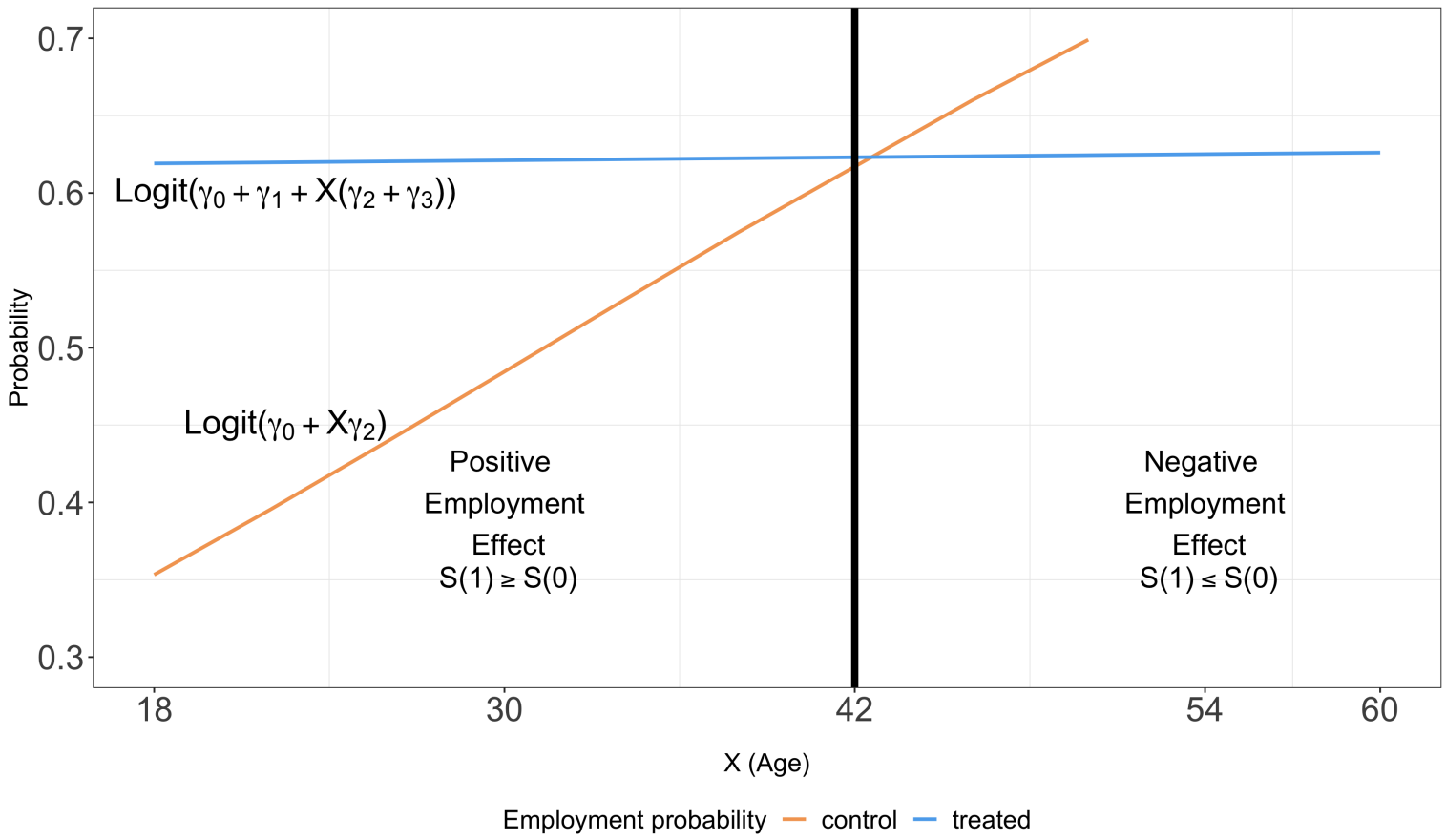
$$\widehat{\beta}_{\text{better}} = 0.90 > \widehat{\beta}_{\text{discrete}} = 0.85$$

Better Lee Bounds: conditional monotonicity

employment probability

$$s(D, X) = \Pr(S = 1 \mid D, X) = \text{Logit}(\gamma_0 + D \cdot \gamma_1 + X \cdot \gamma_2 + D \cdot X \cdot \gamma_3)$$

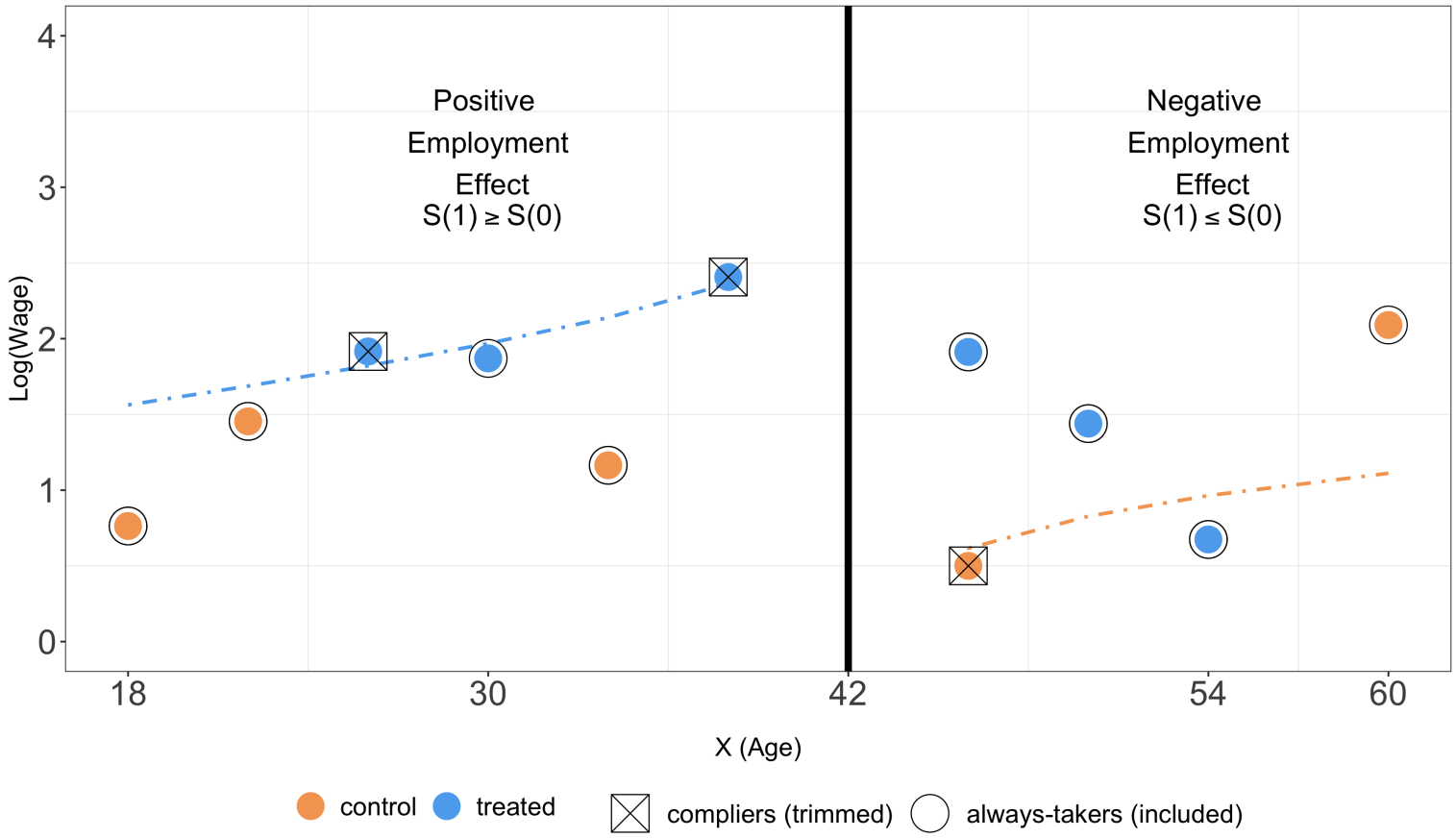
$$X_{\text{pos}} = \{X : s(0, X) < s(1, X)\} \text{ and } X_{\text{neg}} = \{X : s(0, X) > s(1, X)\}$$



Better Lee Bounds: conditional monotonicity

the always-takers' share is $p(X)$ if $s(0,X) < s(1,X)$ and $1/p(X)$ otherwise

the borderline wage is $\begin{cases} \text{quant}_1(X, p(X)) & s(0,X) < s(1,X) \\ \text{quant}_0(X, 1 - 1/p(X)) & s(0,X) > s(1,X) \end{cases}$



Better Lee Bounds: data-driven covariate choice

1. wage and employment equations are sparse: bounds are asymptotically sharp
   - 1st stage: logistic and quantile LASSO + automated penalty choice as in (Belloni et al (ECMA, 2017))
   - 2nd stage: bias correction (Double Machine Learning, Chernozhukov et al. (2018))
2. wage equation is not sparse
   - bounds width is proportional to first-stage $R^2 \Rightarrow$ rank covariates by unexplained variance!
   - 1st stage: logistic and quantile series
   - sharpness is not guaranteed, but tighter in practice

Outline

1. Overview of Lee (2009) bounds
2. Better Lee Bounds
3. JobCorps revisited

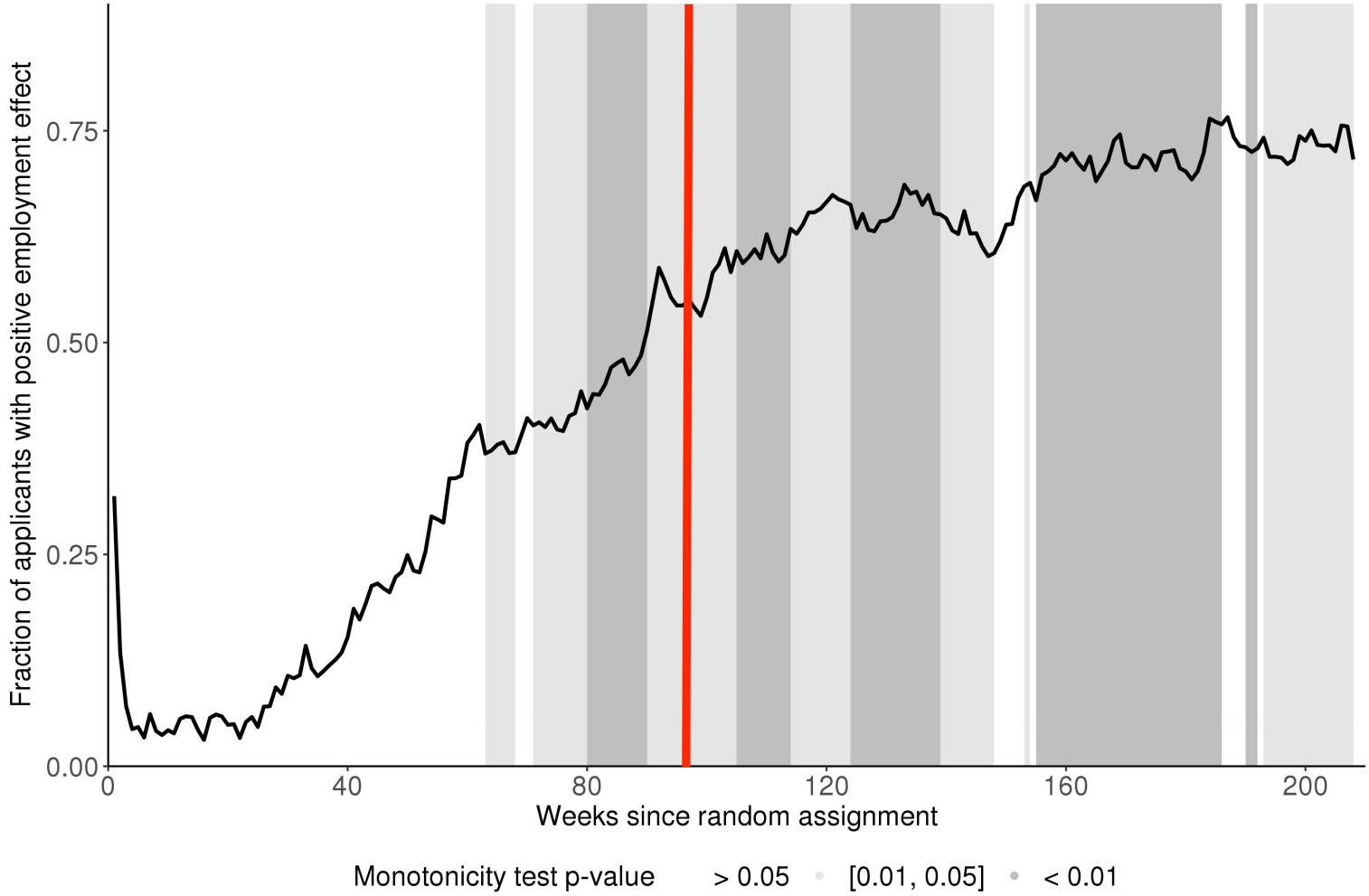
JobCorps: overview of the program

- JobCorps is a training program that helps youth ages 16 through 24 to get a better job, make more money, and gain control over their lives (U.S. Department of Labor, 2005b).

JobCorps: overview of the program

- JobCorps is a training program that helps youth ages 16 through 24 to get a better job, make more money, and gain control over their lives (U.S. Department of Labor, 2005b).
- $N = 9,145$ applicants – the sample sample as in Lee (2009)
- $p = 5,177$ baseline covariates $=$ demographics, reasons for joining JobCorps, medical, arrest, and drug use records, wage history
  - Lee (2009): 28 demographic covariates
- Lee (2009): week 90 wage effect is $[0.048, 0.049]$

JobCorps: monotonicity failure demonstration



week 90: JobCorps helps (hurts) employment for 50 % of subjects

Lee's week 90 estimates $= [0.048, 0.049]$

JobCorps: bounds on week 90 wage effect

	(1) series	(2) lasso	(3) union	(4) agnostic
est. bounds	[-0.005, 0.091]	[0.040, 0.046]	[0.041, 0.059]	[0.041, 0.043]
95 % CI	(-0.05, 0.135)	(0.001, 0.078)	(-0.02, 0.112)	(-0.02, 0.101)
# emplmnt covs	28	5 177 (9)	15	12
# wage covs	28	421 (6)	15	12

1. 28 Lee's covs
2. 5, 177 = all covs, (9) = emplmnt equation, (6) = wage equation
3. 15 = union of employment and wage covs selected in Column 2
4. 12 covs selected on the sample Lee excluded

JobCorps: bounds on week 90 wage effect

	(1) series	(2) lasso	(3) union	(4) agnostic
est. bounds	[-0.005, 0.091]	[0.040, 0.046]	[0.041, 0.059]	[0.041, 0.043]
95 % CI	(-0.05, 0.135)	(0.001, 0.078)	(-0.02, 0.112)	(-0.02, 0.101)
# emplmnt covs	28	5 177 (9)	15	12
# wage covs	28	421 (6)	15	12

1^{st} **stage** :

28 Lee's covs = age, race, education, parental educ., income, earnings at baseline

logit Employed D $(X_1 - X_{28})$ $D * (X_1 - X_{28})$

qreg LogWage D $(X_1 - X_{28})$ $D * (X_1 - X_{28})$ if Employed $== 1$

JobCorps: bounds on week 90 wage effect

	(1) series	(2) lasso	(3) union	(4) agnostic
est. bounds	[-0.005, 0.091]	[0.040, 0.046]	[0.041, 0.059]	[0.041, 0.043]
95 % CI	(-0.05, 0.135)	(0.001, 0.078)	(-0.02, 0.112)	(-0.02, 0.101)
# emplmnt covs	28	5 177 (9)	15	12
# wage covs	28	421 (6)	15	12

1^{st} **stage** :

5177=28 covs + reasons for joining JobCorps, medical, arrest, and drug use records, wage history

post lasso logit Employed D $(X_1 - X_{5177})$ $D * (X_1 - X_{5177})$

post lasso qreg LogWage D $(X_1 - X_{421})$ $D * (X_1 - X_{421})$ if Employed $== 1$

JobCorps: bounds on week 90 wage effect

	(1) series	(2) lasso	(3) union	(4) agnostic
est. bounds	[-0.005, 0.091]	[0.040, 0.046]	[0.041, 0.059]	[0.041, 0.043]
95 % CI	(-0.05, 0.135)	(0.001, 0.078)	(-0.02, 0.112)	(-0.02, 0.101)
# emplmnt covs	28	5 177 (9)	15	12
# wage covs	28	421 (6)	15	12

1^{st} **stage** :

$$X = 15 = 9 \text{ emplmnt} + 6 \text{ wage covs}$$

logit Employed D $(X_1 - X_{15})$ $D * (X_1 - X_{15})$

qreg LogWage D $(X_1 - X_{15})$ $D * (X_1 - X_{15})$ if Employed $== 1$

JobCorps: bounds on week 90 wage effect

	(1) series	(2) lasso	(3) union	(4) agnostic
est. bounds	[-0.005, 0.091]	[0.040, 0.046]	[0.041, 0.059]	[0.041, 0.043]
95 % CI	(-0.05, 0.135)	(0.001, 0.078)	(-0.02, 0.112)	(-0.02, 0.101)
# emplmnt covs	28	5 177 (9)	15	12
# wage covs	28	421 (6)	15	12

1^{st} **stage** :

12 covs explaining most variation in wage

logit Employed D $(X_1 - X_{12})$ $D * (X_1 - X_{12})$

qreg LogWage D $(X_1 - X_{12})$ $D * (X_1 - X_{12})$ if Employed $== 1$

Conclusion

Today:

- ✓ a smooth-based solution to degenerate cell problem
- ✓ discovered composition effects in JobCorps

Conclusion

Today:

- ✓ a smooth-based solution to degenerate cell problem
- ✓ discovered composition effects in JobCorps

This paper and `leebounds` R package:

- \+ non-compliance
- \+ unknown propensity score (e.g., quasi-experiments)

Conclusion

Today:

- ✓ a smooth-based solution to degenerate cell problem
- ✓ discovered composition effects in JobCorps

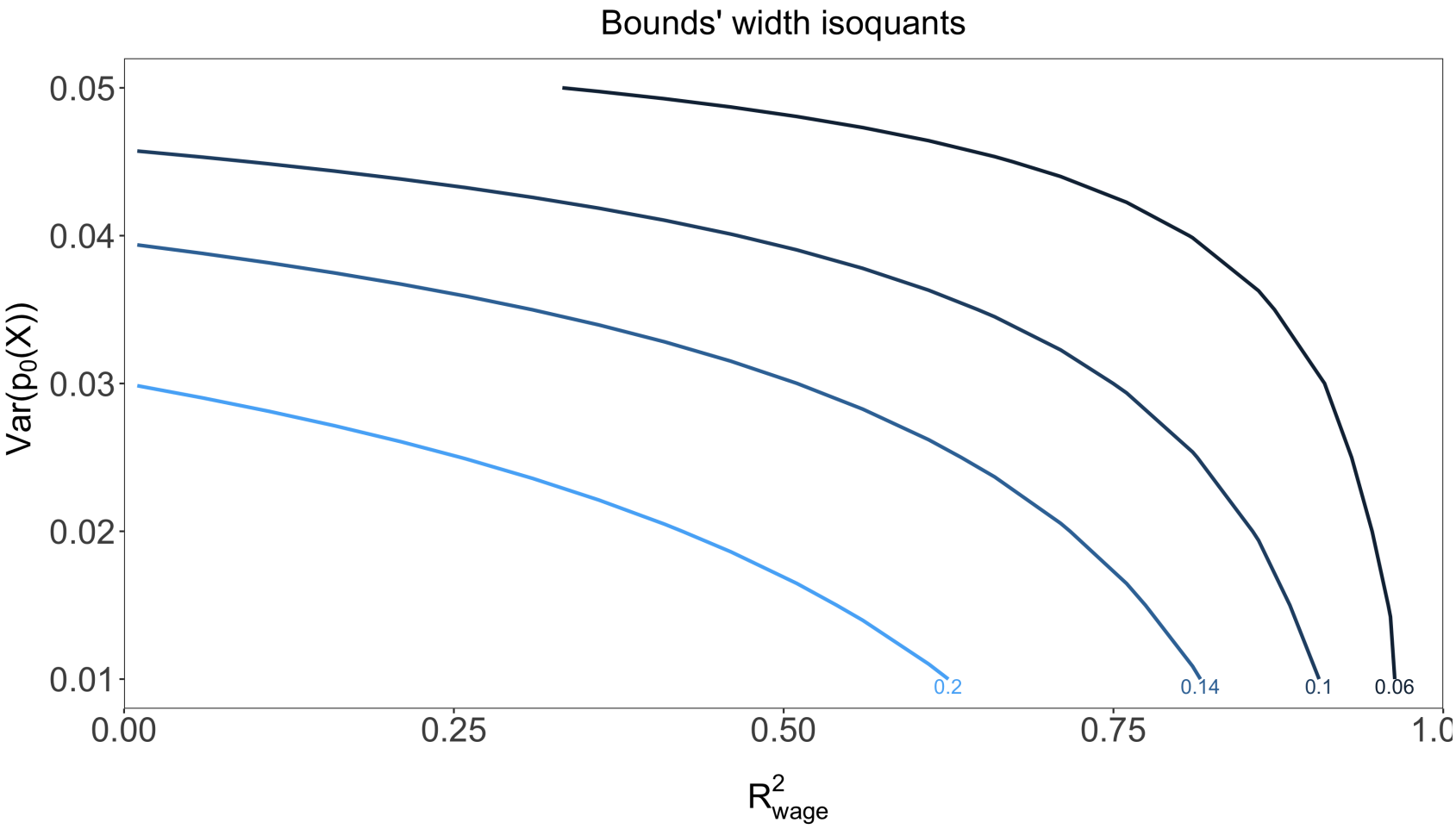
This paper and `leebounds` R package:

- \+ non-compliance
- \+ unknown propensity score (e.g., quasi-experiments)

Future work

- sharp bounds $\neq$ tightest CI!
- strong partial ID $<$ very tight bounds $<$ point ID

Intuition for bounds' width



Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal*, 21:C1–C68.

Kolesar, M. (2013). Estimation in an instrumental variable model with treatment effect heterogeneity.

Lee, D. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3):1071–1102.