The Effect of Maternal Labor Supply on Children: Evidence from Bunching

Carolina Caetano, Gregorio Caetano,
Eric Nielsen and Viviane Sanfelice

UGA/Fed Board/Temple Univ.

Disclaimer: The views and opinions expressed in this presentation are solely those of the authors and should not be interpreted as reflecting the official policy or position of the Board of Governors or the Federal Reserve System.

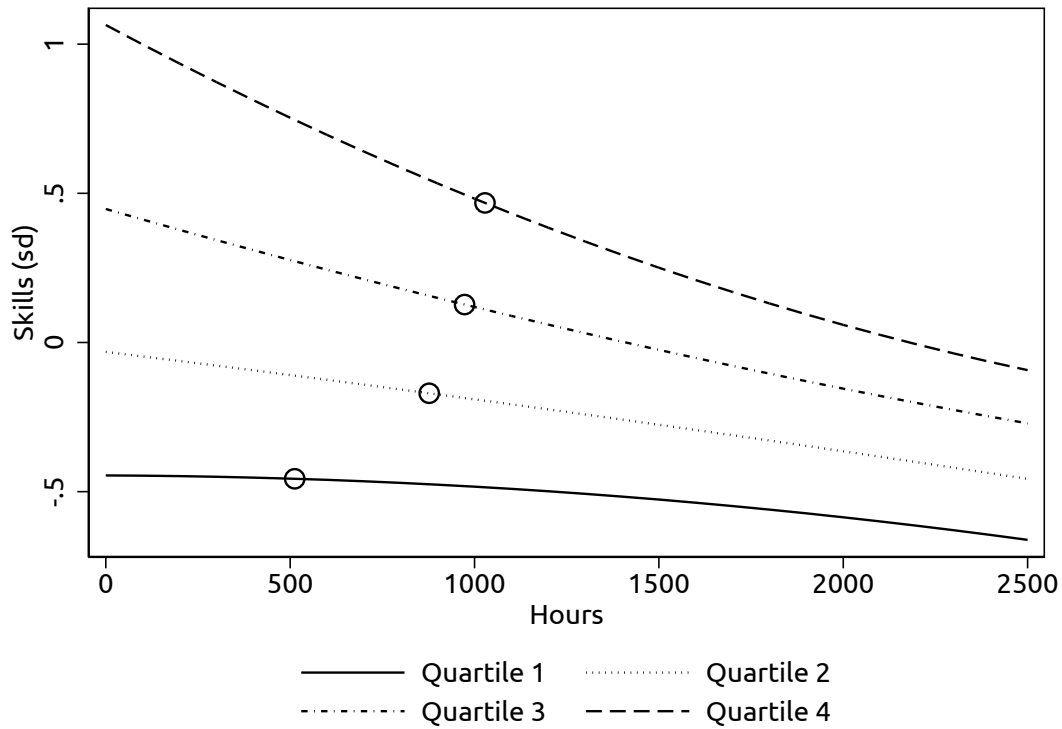
Motivation

- Maternal labor supply has increased in recent decades (Eckstein and Lifschitz 2011, Fogli and Veldkamp 2011).
- Quality parent-child interactions known to be important for child development (e.g. Todd and Wolpin 2007).
- Effect of maternal labor supply on children's skills in the short-run
  1. time channel: more time at work $\Longrightarrow$ less time at home
  2. income channel: more time at work $\Longrightarrow$ more income
- Many policies affect maternal labor supply (family leave, paid childcare, child tax credits, etc.).
- Many other reasons for work-promoting policies.

This paper

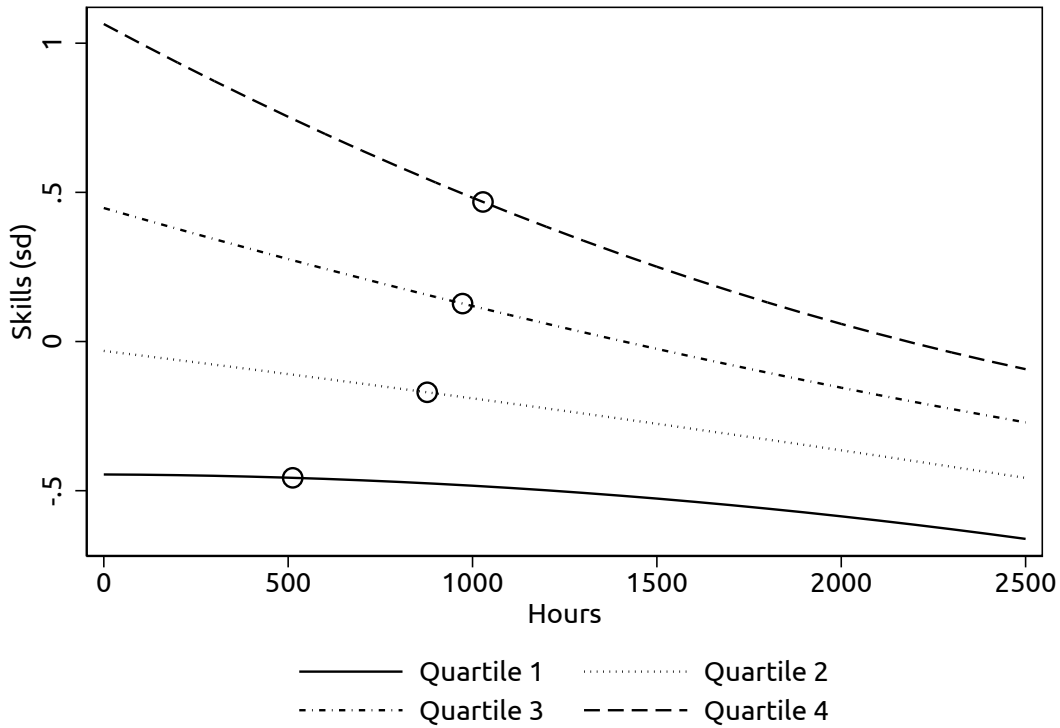
- Effect of mothers working hours during first 3 years of child's life on child's skill around age 6.
- A new approach to deal with endogeneity in this context (Caetano, Caetano and Nielsen 2021).
- We focus on heterogeneous effects:
  - By the skill of the mother.
  - By the quantity of her labor supply.

Preview of Results: By quartile of mother AFQT



- Effect tends to be negative for children's cognitive skills in the short-run.
- Less negative in the case of low-skilled mothers, except those who work long hours.

Preview of Results: By quartile of mother AFQT



- Why are higher skilled mothers' labor supply so detrimental for children's skills in the short-run?
  - ~~Because last hour is more costly for those working longer hours.~~
  - Money is not enough to compensate in the short-run for their high-skilled interaction with the child.

A third source of heterogeneity: By pre-birth wage rate

- Think of two high-skilled mothers, one with an economics major, another with a language major.
- Try to vary income holding quality of home interaction constant. The job-related skills that separate them are well valued in the job market, but not as much in the interaction with a 0-3 year old.
- Money helps, but is not enough: Even the high-skilled, high-paid mother cannot fully compensate for her absence with money.
- We interpret that giving flexibility to mothers to work from home is likely a better policy than giving financial incentives to work.

Literature

- NLSY79/CNLSY. Effect of working hours while child is 0-3 on child's early outcomes.
  - Negative effects:

    Ruhm 2004, Bernal 2008, Desai et al. 1989, Baydar and Brooks-Gunn 1991, Hill and O'Neill 1994, Baum 2003
  - Zero effects:

    James-Burdumy 2005, Parcel and Menaghan 1994, Blau et al. 1992, Waldfogel et al. 2002
  - Positive effects:

    Vandell and Ramanan 1992 (low-income families)
- More recent literature: focus on time-money trade-off
  - Negative net effects:

    Agostinelli and Sorrenti (2021): NLSY79/CNLSY. Effect of working hours while child is 4-16 on contemporaneous child outcome.
  - Zero net effects:

    Nicoletti, Salvanes and Tominey (2022): Norwegian data. Effect of working hours while child is 0-5 on outcome when child is 11-15.

Data

NLSY79 (Mothers)

- AFQT, education, marital status, age at birth, hh income
- average hours worked over three years following birth
- exclude three months immediately after birth (maternity leave)
- family structure

CNLSY (Children)

- race/ethnicity, sex
- cognitive skills – iterated principal factor analysis applied to PIAT math and reading scores, administered around age 6

Final sample consists of 6,924 mother-child pairs.

Basic set up (no controls)

$$S = \alpha + \beta L + \delta L^* + \epsilon$$

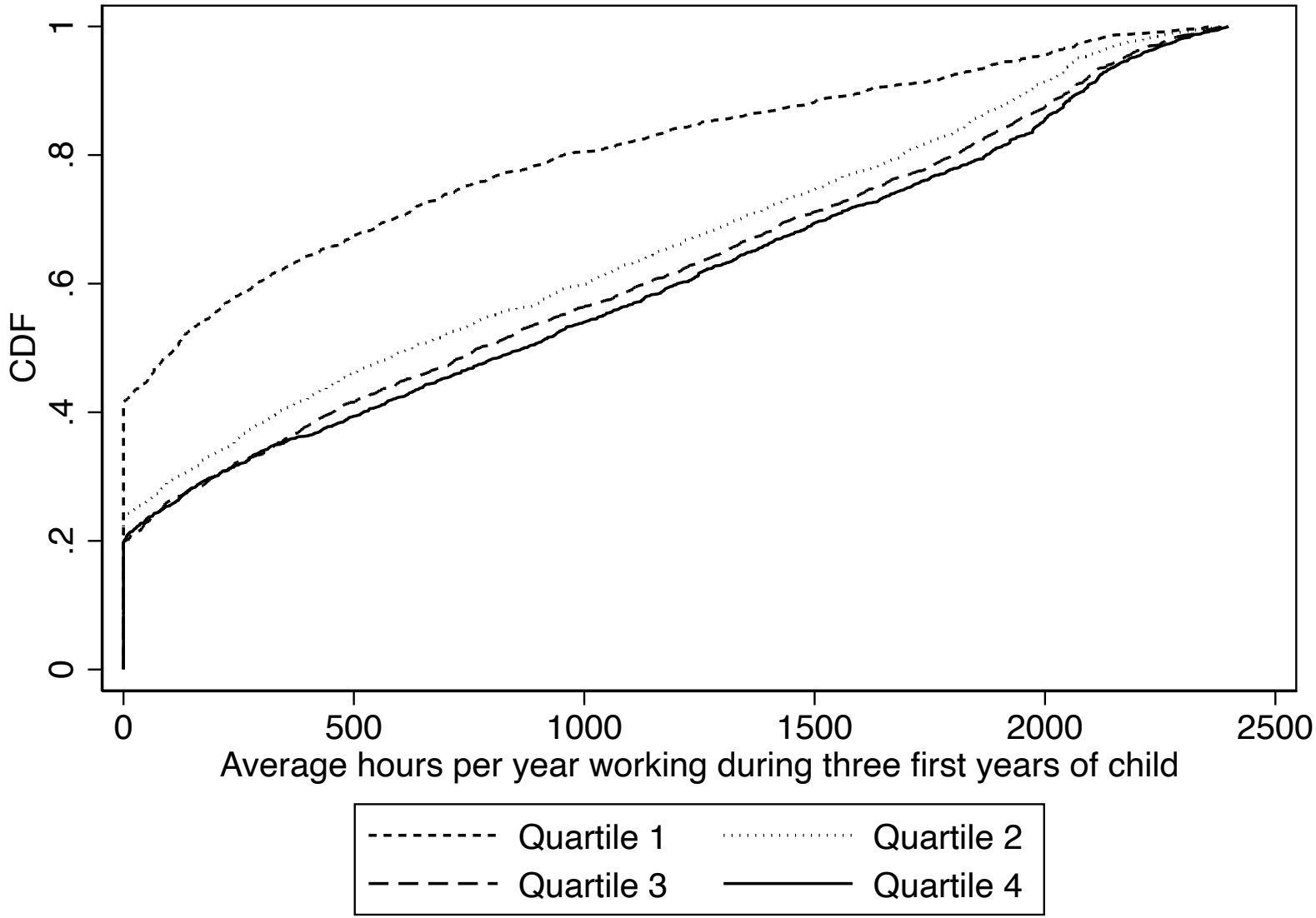
What is the effect of increasing hours from $L = l_0$ to $L = l_1$?

$$\underbrace{\mathbb{E}[S|L = l_1, L^* = l_1] - \mathbb{E}[S|L = l_0, L^* = l_0]}_{\text{what we observe}} =$$

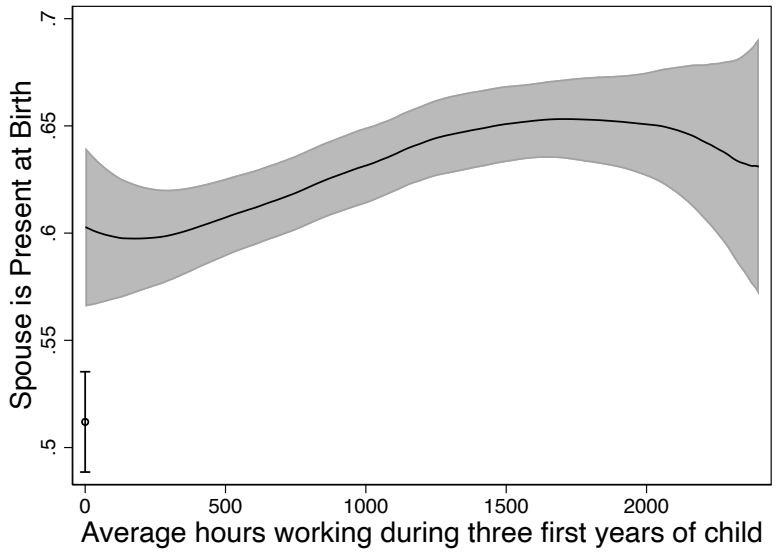
$$\underbrace{\mathbb{E}[S|L = l_1, L^* = l_1] - \mathbb{E}[S|L = l_0, L^* = l_1]}_{\text{marginal treatment effect}} +$$

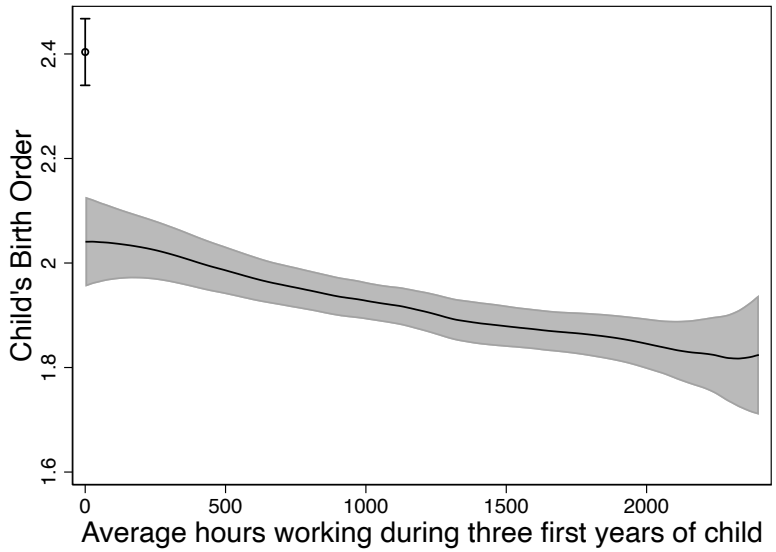
$$\underbrace{\mathbb{E}[S|L = l_0, L^* = l_1] - \mathbb{E}[S|L = l_0, L^* = l_0]}_{\text{selection bias}}$$

Bunching of Labor Supply by Quartile of AFQT score

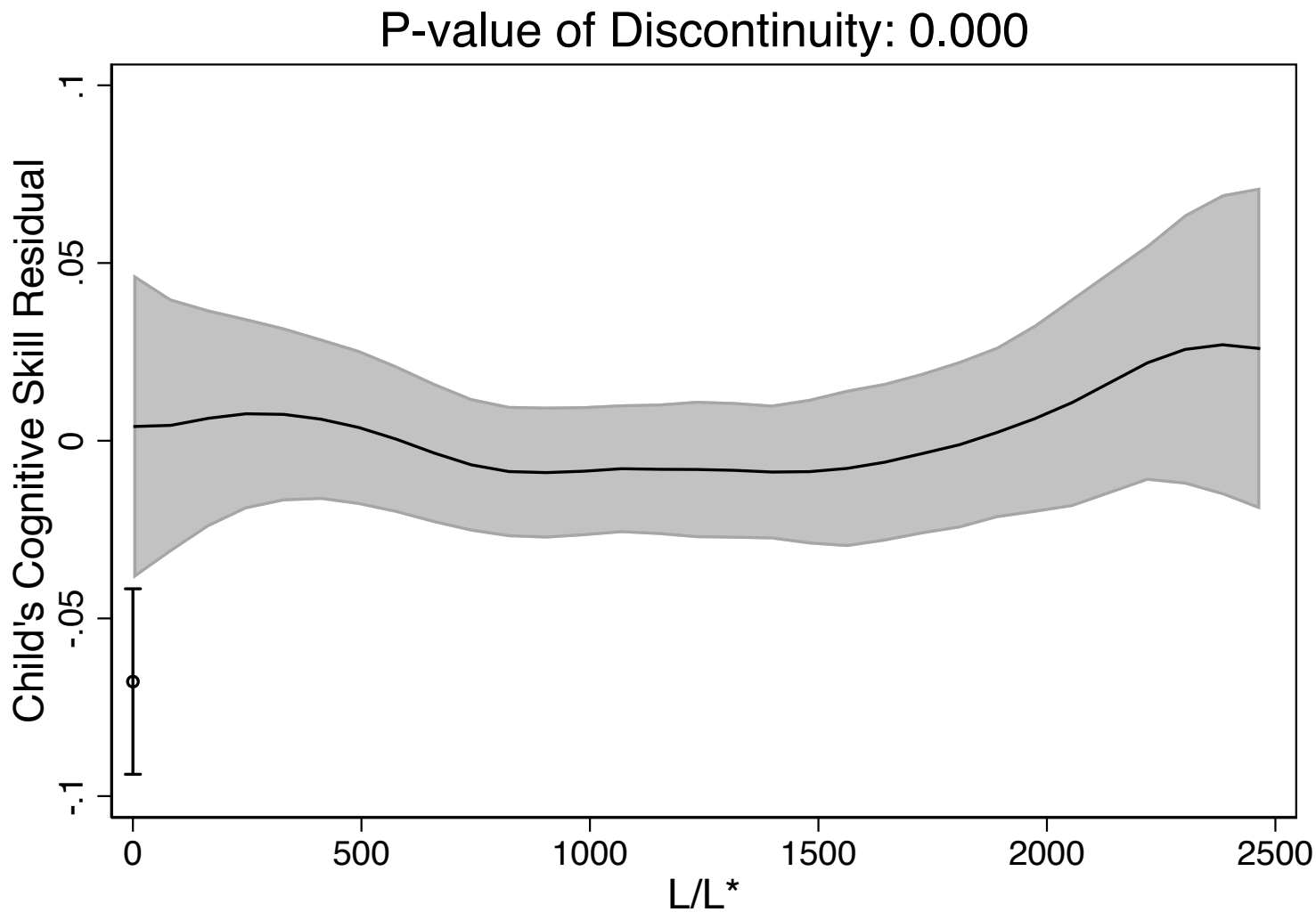


Selective Bunching





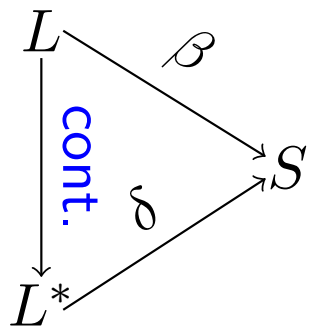
Are observed controls enough?

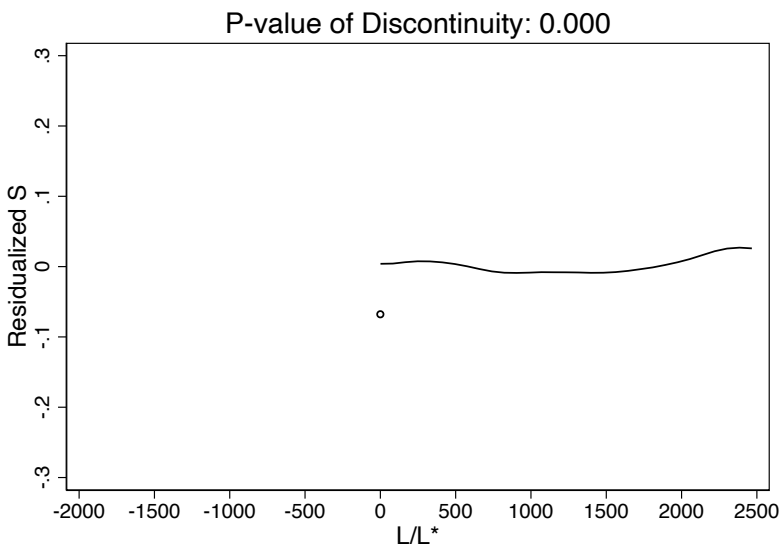


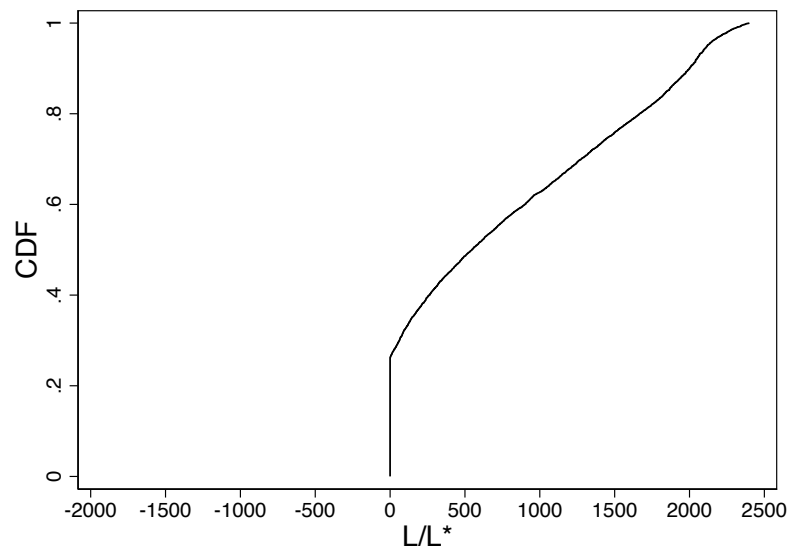
Intuition

▸ RDD ▸ Linearity ▸ Formal Model w/ Covariates

$$S = \alpha + \beta L + \delta L^* + \epsilon, \qquad L = \max\{0, L^*\}$$



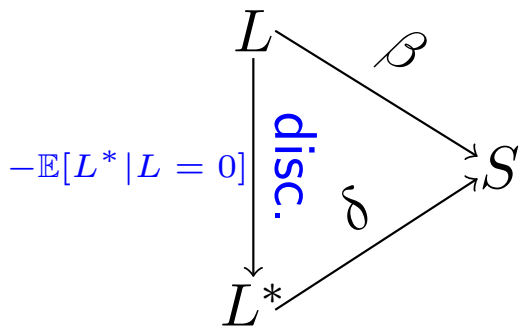


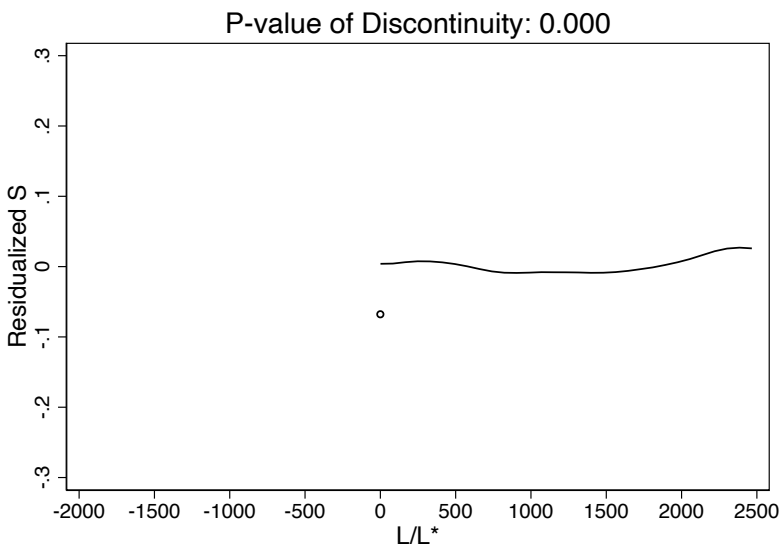


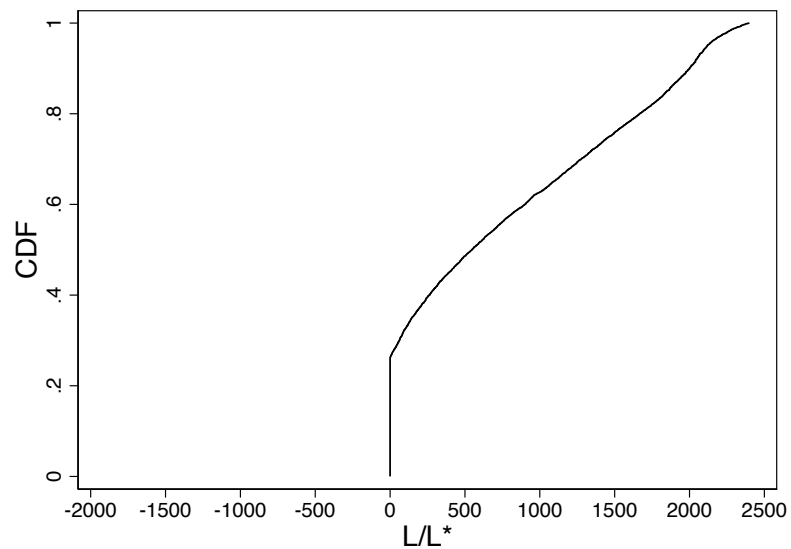
Intuition

▸ RDD ▸ Linearity ▸ Formal Model w/ Covariates

$$S = \alpha + \beta L + \delta L^* + \epsilon, \qquad L = \max\{0, L^*\}$$



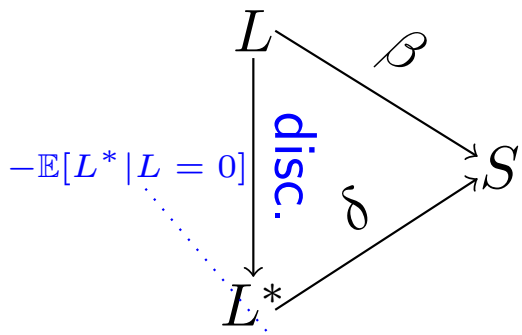


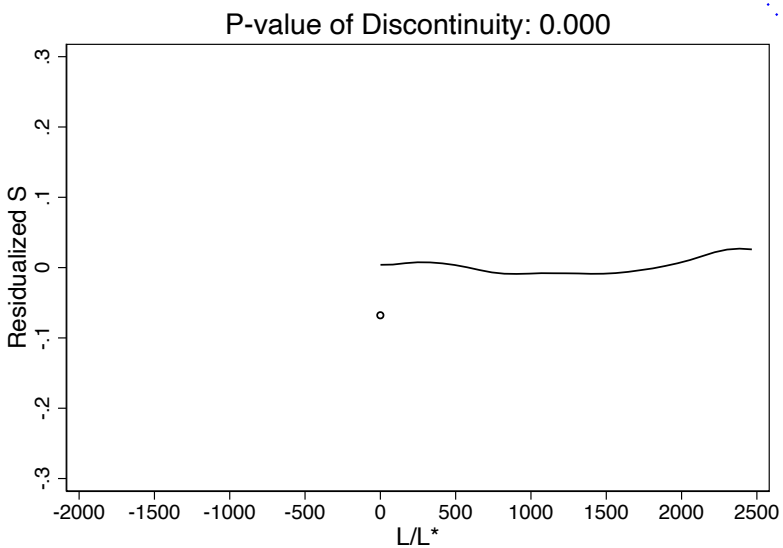


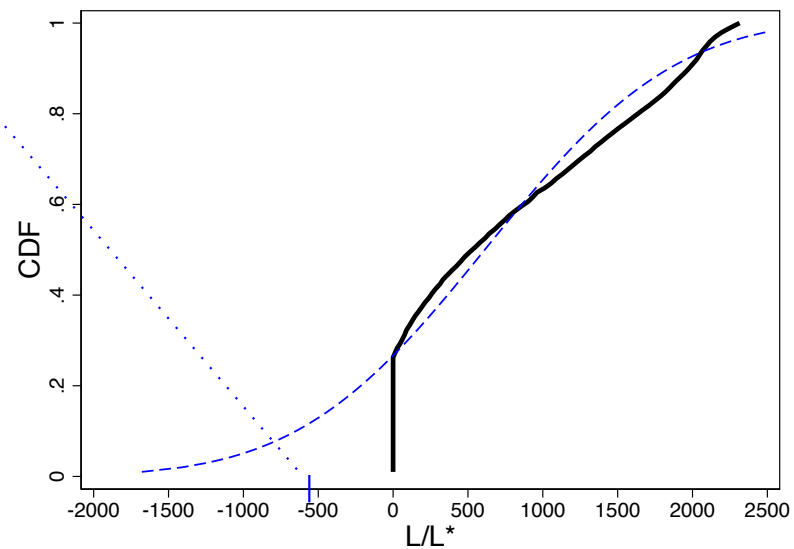
Intuition

▸ RDD ▸ Linearity ▸ Formal Model w/ Covariates

$$S = \alpha + \beta L + \delta L^* + \epsilon, \qquad L = \max\{0, L^*\}$$



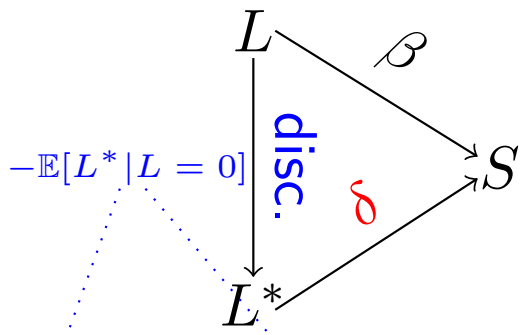


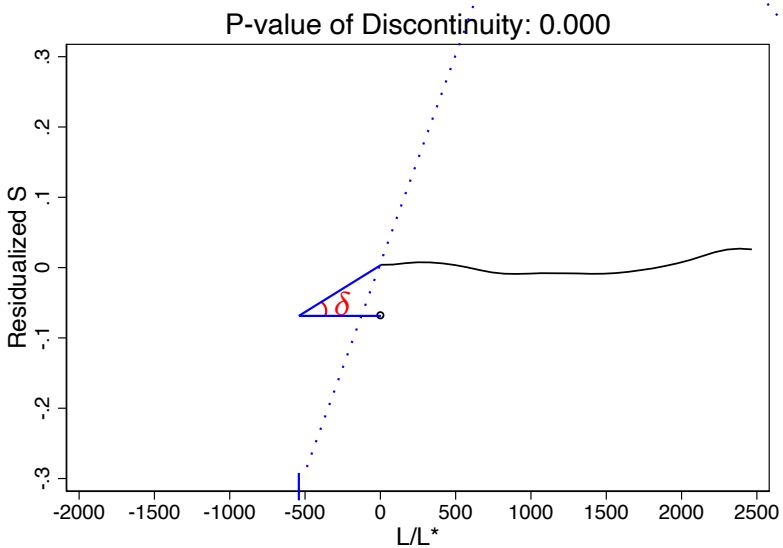


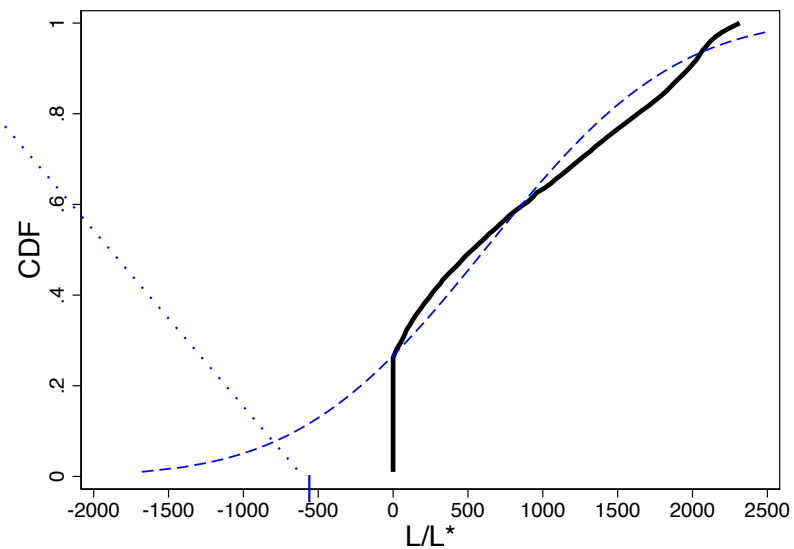
Intuition

▸ RDD ▸ Linearity ▸ Formal Model w/ Covariates

$$S = \alpha + \beta L + \delta L^* + \epsilon, \qquad L = \max\{0, L^*\}$$



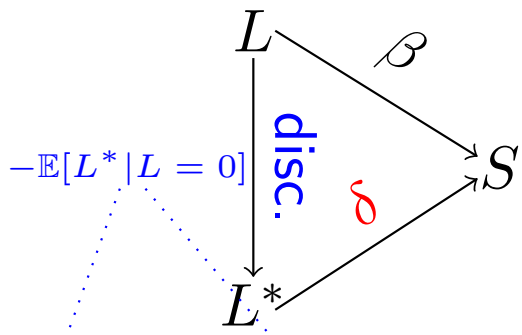


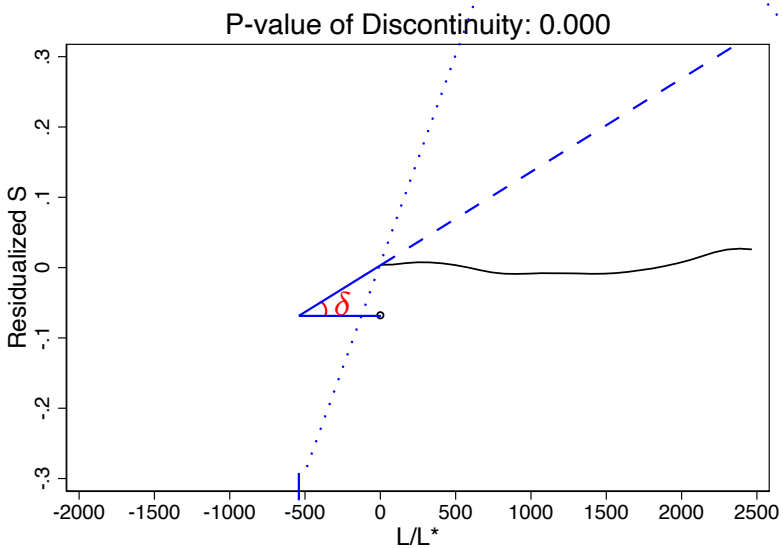


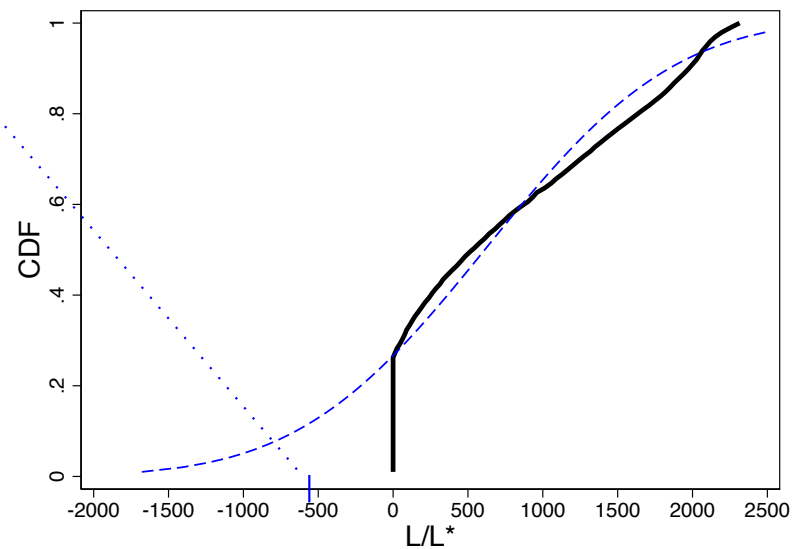
Intuition

▸ RDD ▸ Linearity ▸ Formal Model w/ Covariates

$$S = \alpha + \beta L + \delta L^* + \epsilon, \qquad L = \max\{0, L^*\}$$



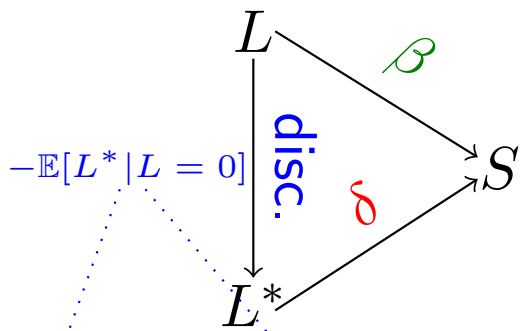


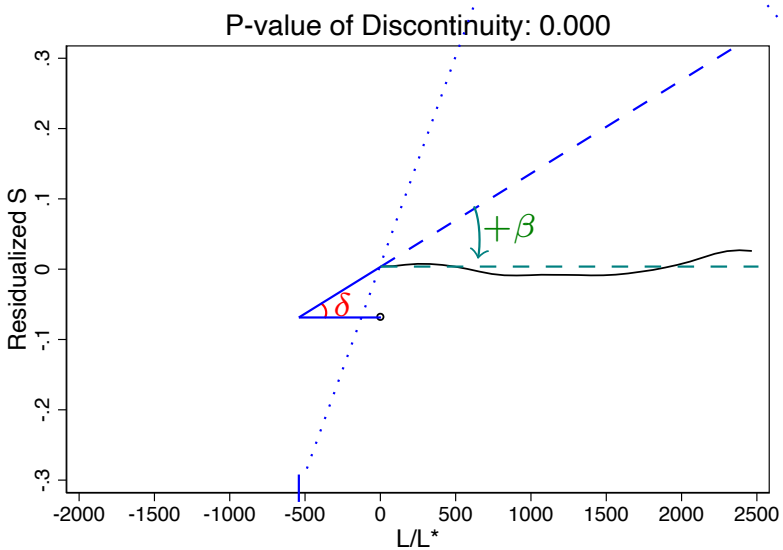


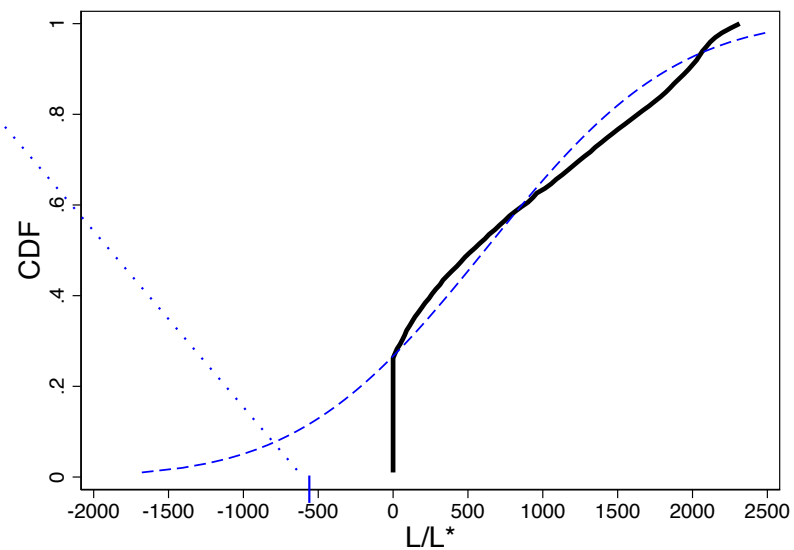
Intuition

▸ RDD ▸ Linearity ▸ Formal Model w/ Covariates

$$S = \alpha + \beta L + \delta L^* + \epsilon, \qquad L = \max\{0, L^*\}$$







$$S = \beta L + g(X) + \delta\eta + \varepsilon, \quad \eta = L^* - h(X)$$

	(i) No Controls	(ii) Controls	(iii) Het. Tobit	(iv) Het. Uniform	(v) Het. Symmetric
β	0.014**	0.000	-0.016**	-0.019**	**-0.019****
	(0.001)	(0.001)	(0.005)	(0.006)	**(0.005)**
δ			0.014**	0.017**	**0.017****
			(0.004)	(0.005)	**(0.005)**

Note: This table shows estimates of the effect of an additional 100 hours per year working in the three years following the child's birth on the child's early cognitive skills. N=6,924. Bootstrapped standard errors in parentheses (1,000 bootstrap samples). ** $p<0.05$, * $p<0.1$.

Heterogeneity by Maternal Skill and Hours

$$S = \beta(L, X) + g(X) + \delta(X)\eta + \varepsilon$$

$A =$ AFQT score of the mother

$$\beta(L, X) = \beta L + \beta_A AL + \beta_L L^2 + \beta_{AL} AL^2$$

$$\delta(X) = \delta + \delta_A A$$

Heterogeneity by Maternal Skill and Hours

	(i) No Controls	(ii) Controls	(iii) Het. Tobit	(iv) Het. Uniform	(v) Het. Symmetric
β	0.018**	0.003	-0.023**	-0.026**	**-0.030****
	(0.004)	(0.003)	(0.008)	(0.010)	**(0.009)**
β_A	0.047**	-0.010**	-0.017**	-0.021**	**-0.026****
	(0.003)	(0.004)	(0.008)	(0.009)	**(0.009)**
β_L ($\times$1000)	-0.004**	-0.001	0.002	0.001	**0.002**
	(0.002)	(0.001)	(0.002)	(0.002)	**(0.002)**
β_{AL} ($\times$1000)	-0.015**	0.003**	0.004**	0.004**	**0.004****
	(0.001)	(0.002)	(0.002)	(0.002)	**(0.002)**
δ			0.016**	0.022**	**0.023****
			(0.005)	(0.007)	**(0.006)**
δ_A			0.005	0.009	**0.013****
			(0.005)	(0.006)	**(0.006)**

Note: This table shows estimates of the effect of an additional 100 hours per year working in the three years following the child's birth on the child's early cognitive skills. N=6,924. Bootstrapped standard errors in parentheses (1,000 bootstrap samples). The corrected specifications use $K = 50$ clusters and include cluster indicators as controls. ** p<0.05, * p<0.1.

Heterogeneity by Maternal Skill and Hours

	(i) No Controls	(ii) Controls	(iii) Het. Tobit	(iv) Het. Uniform	(v) Het. Symmetric
β	0.018**	0.003	-0.023**	-0.026**	**-0.030****
	(0.004)	(0.003)	(0.008)	(0.010)	**(0.009)**
β_A	0.047**	-0.010**	-0.017**	-0.021**	**-0.026****
	(0.003)	(0.004)	(0.008)	(0.009)	**(0.009)**
β_L ($\times$1000)	-0.004**	-0.001	0.002	0.001	**0.002**
	(0.002)	(0.001)	(0.002)	(0.002)	**(0.002)**
β_{AL} ($\times$1000)	-0.015**	0.003**	0.004**	0.004**	**0.004****
	(0.001)	(0.002)	(0.002)	(0.002)	**(0.002)**
δ			0.016**	0.022**	**0.023****
			(0.005)	(0.007)	**(0.006)**
δ_A			0.005	0.009	**0.013****
			(0.005)	(0.006)	**(0.006)**

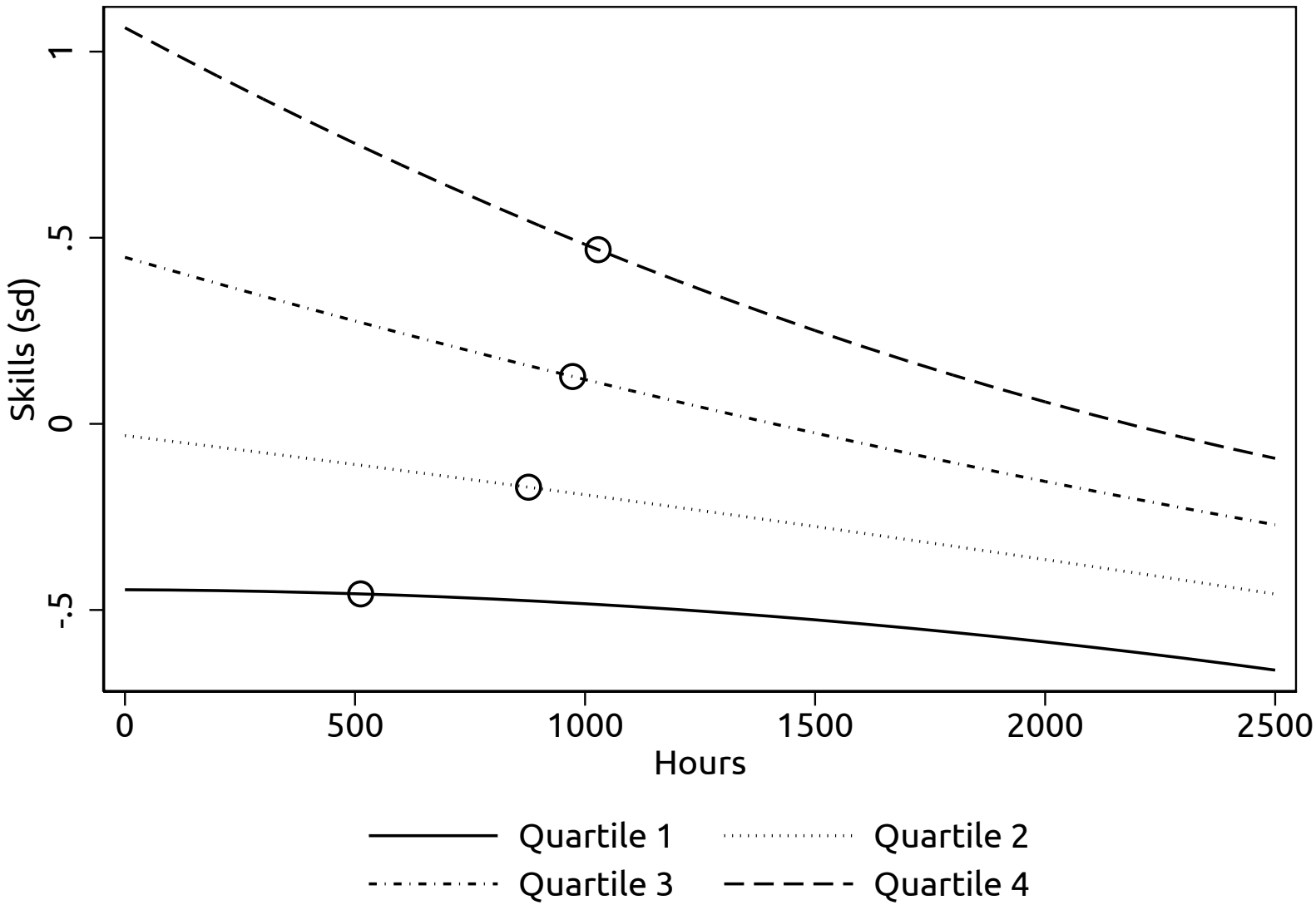
Note: This table shows estimates of the effect of an additional 100 hours per year working in the three years following the child's birth on the child's early cognitive skills. N=6,924. Bootstrapped standard errors in parentheses (1,000 bootstrap samples). The corrected specifications use $K = 50$ clusters and include cluster indicators as controls. ** p<0.05, * p<0.1.

Heterogeneity by Maternal Skill and Hours

	(i) No Controls	(ii) Controls	(iii) Het. Tobit	(iv) Het. Uniform	(v) Het. Symmetric
β	0.018**	0.003	-0.023**	-0.026**	**-0.030****
	(0.004)	(0.003)	(0.008)	(0.010)	**(0.009)**
β_A	0.047**	-0.010**	-0.017**	-0.021**	**-0.026****
	(0.003)	(0.004)	(0.008)	(0.009)	**(0.009)**
β_L ($\times$1000)	-0.004**	-0.001	0.002	0.001	**0.002**
	(0.002)	(0.001)	(0.002)	(0.002)	**(0.002)**
β_{AL} ($\times$1000)	-0.015**	0.003**	0.004**	0.004**	**0.004****
	(0.001)	(0.002)	(0.002)	(0.002)	**(0.002)**
δ			0.016**	0.022**	**0.023****
			(0.005)	(0.007)	**(0.006)**
δ_A			0.005	0.009	**0.013****
			(0.005)	(0.006)	**(0.006)**

Note: This table shows estimates of the effect of an additional 100 hours per year working in the three years following the child's birth on the child's early cognitive skills. N=6,924. Bootstrapped standard errors in parentheses (1,000 bootstrap samples). The corrected specifications use $K = 50$ clusters and include cluster indicators as controls. ** p<0.05, * p<0.1.

Visualizing the Heterogeneous Effects



Further Heterogeneity: by Wage

$$S = \beta(L, X) + g(X) + \delta(X)\eta + \varepsilon$$

$A =$ AFQT score, $W =$ Pre-birth wage

$$\beta(L, X) = \beta L + \beta_A A L + \beta_W W L + \beta_L L^2 + \beta_{AL} A L^2 + \beta_{WL} W L^2$$

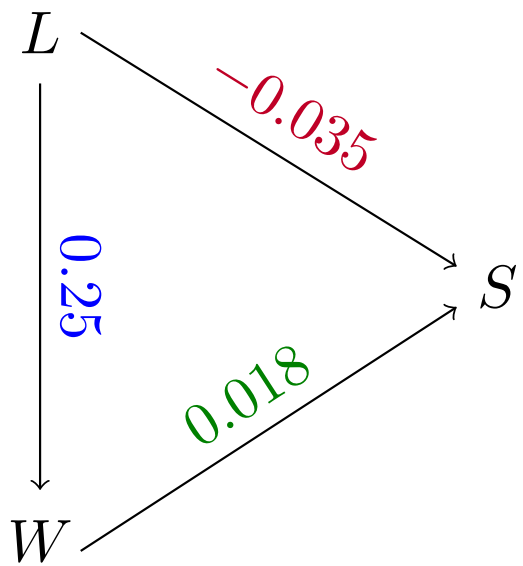
$$\delta(X) = \delta + \delta_A A + \delta_W W$$

Heterogeneity by Maternal Skill, Wages, and Hours

	(i) No Controls	(ii) Controls	(iii) Het. Tobit	(iv) Het. Uniform	(v) Het. Symmetric
β	-0.004	-0.002	-0.009	-0.013	**-0.017**
	(0.005)	(0.004)	(0.014)	(0.026)	**(0.017)**
β_A	0.039**	-0.016**	-0.027**	-0.047**	**-0.035****
	(0.004)	(0.004)	(0.014)	(0.023)	**(0.017)**
β_W	0.007*	0.001	0.014	0.029	**0.018**
	(0.004)	(0.005)	(0.016)	(0.026)	**(0.019)**
β_L ($\times$1000)	0.002	0.001	0.001	0.001	**0.001**
	(0.002)	(0.002)	(0.002)	(0.002)	**(0.002)**
β_{AL} ($\times$1000)	-0.012**	0.005**	0.006**	0.007**	**0.007****
	(0.002)	(0.002)	(0.002)	(0.002)	**(0.002)**
β_{WL} ($\times$1000)	-0.002	0.001	-0.000	-0.001	**-0.000**
	(0.002)	(0.002)	(0.002)	(0.002)	**(0.002)**
δ			0.006	0.010	**0.013**
			(0.010)	(0.022)	**(0.013)**
δ_A			0.008	0.027	**0.015**
			(0.009)	(0.020)	**(0.013)**
δ_W			-0.009	-0.024	**-0.014**
			(0.010)	(0.021)	**(0.014)**

Back-of-the-envelope calculation

Compare the marginal effect of L for two mothers: $A = 1$ and $A = 0$:



$$\beta(A=1) - \beta(A=0) = \underbrace{-0.035}_{\text{OC of time}} + \Big(\underbrace{0.25}_{\text{Return to } A} \times \underbrace{0.018}_{\text{Money}} \Big) = -0.031$$

Conclusion

- Effect of mother working in the first 3 years of the child on child's cognitive skills around age 6: negative on average, very negative for high-skilled mothers.
  - Last hour of work is more costly the longer the mother works?
    No – effects are close to linear and maybe convex for high-skilled mothers.
  - Incremental earnings insufficient to offset direct time effect?
    Yes, even for high-skilled high-wage mothers. Estimates are too noisy though.
- What work-promoting policies would avoid negative effects on children's skills?
  - For low-skilled mothers: no negative effects unless mother works close to full time.
  - Increasing financial rewards to working would likely be ineffective.
  - Flexible work arrangements seem the best bet for higher-skilled mothers.
  - Flexible work for partners may be complementary.

In the paper but not in the presentation

- Formal Discussion of Identification Strategy ▸ Details
- Relation to RDD ▸ Details
- Detailed Sensitivity Analysis ▸ Details

Thanks!

Adding Covariates

Back

$$S = \beta(L, X) + g(X) + \epsilon, \quad \mathbb{E}[\epsilon|L, X] \neq 0$$

Adding Covariates

Back

$$S = \beta(L, X) + g(X) + \overbrace{\delta(X)\eta + \varepsilon}^{\epsilon}, \quad \mathbb{E}[\varepsilon | L, X, \eta] = 0$$

Adding Covariates

Back

$$S = \beta(L, X) + g(X) + \overbrace{\delta(X)\eta + \varepsilon}^{\epsilon}, \quad \mathbb{E}[\varepsilon | L, X, \eta] = 0$$

$$L^* = h(X) + \eta$$

$$L = \max\{0, L^*\}$$

Adding Covariates

Back

$$S = \beta(L, X) + g(X) + \overbrace{\delta(X)\eta + \varepsilon}^{\epsilon}, \quad \mathbb{E}[\varepsilon | L, X, \eta] = 0$$

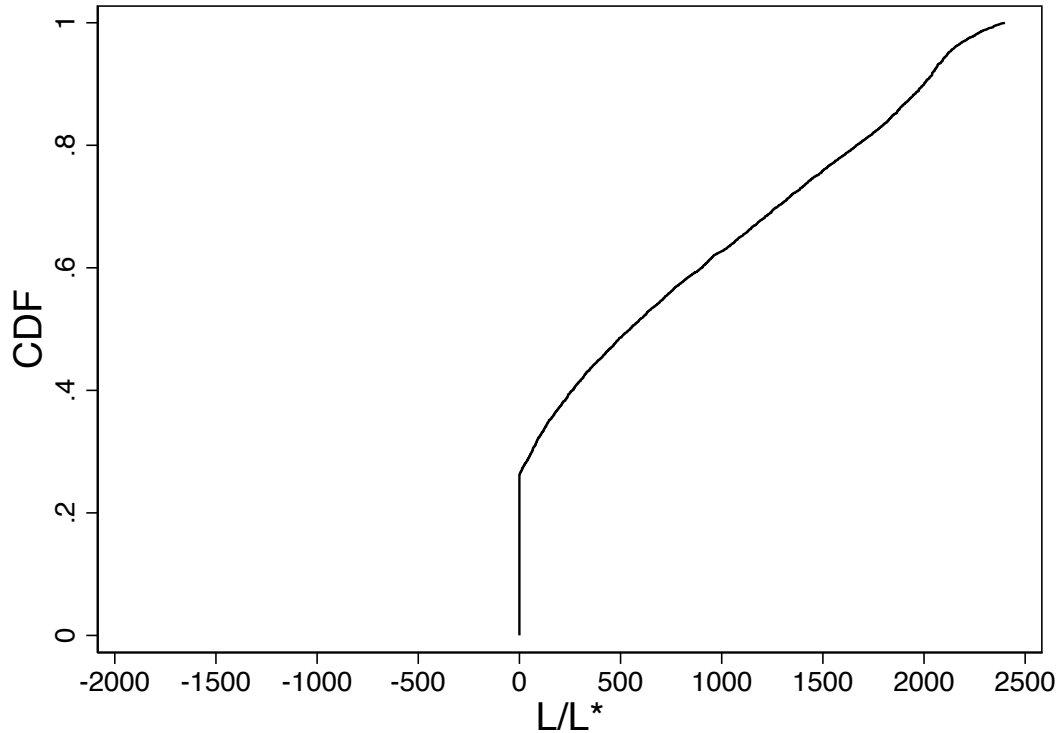
$$L^* = h(X) + \eta$$

$$L = \max\{0, L^*\}$$

$$\mathbb{E}[S|L, X] = \beta(L, X) + \underbrace{g(X) - \delta h(X)}_{m(X)} + \\ + \delta(X) \underbrace{[L + \mathbb{E}[L^*|L = 0, X]1(L = 0)]}_{\text{new regressor}}$$

$\mathbb{E}[L^*|L=0,X]$

▸ Test 1 ▸ Test 2 ▸ Test 3



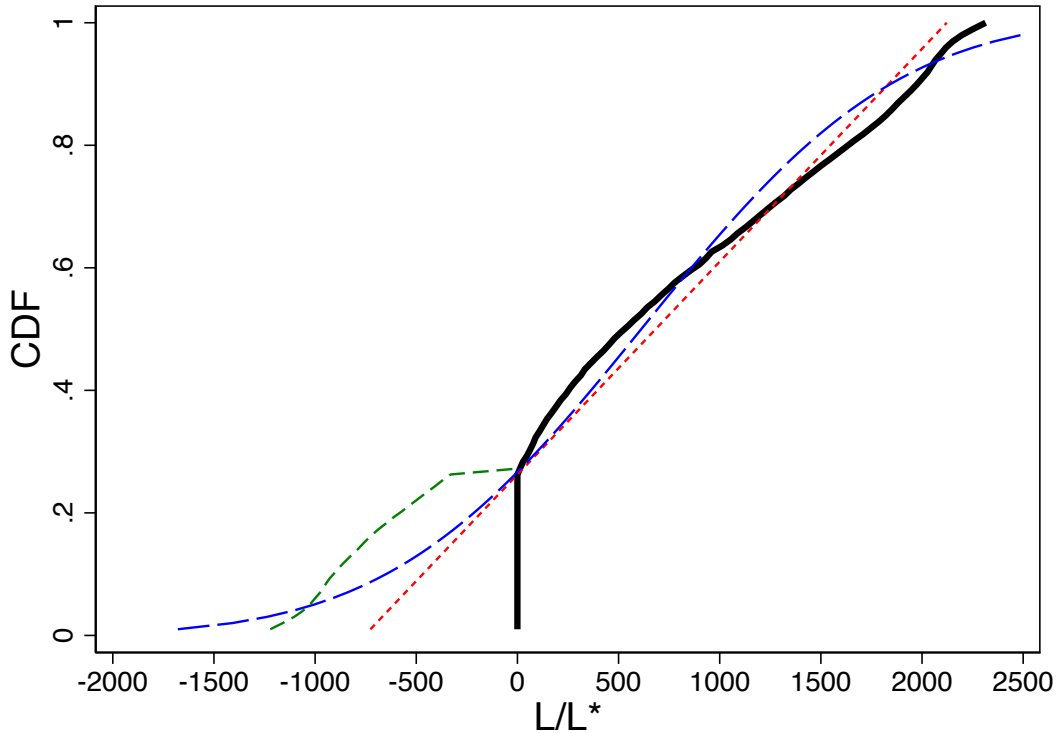
$$S = \beta(L,X) + g(X) + \delta(X)\eta + \varepsilon, \quad \mathbb{E}[\varepsilon|L,X,\eta] = 0$$

$$L^* = h(X) + \eta$$

$$\mathbb{E}[S|L,X] = \beta(L,X) + m(X) + \delta(X)\underbrace{[L + \mathbb{E}[L^*|L=0,X]1(L=0)]}_{\text{new regressor}}$$

$\mathbb{E}[L^*|L=0,X]$

▸ Test 1 ▸ Test 2 ▸ Test 3



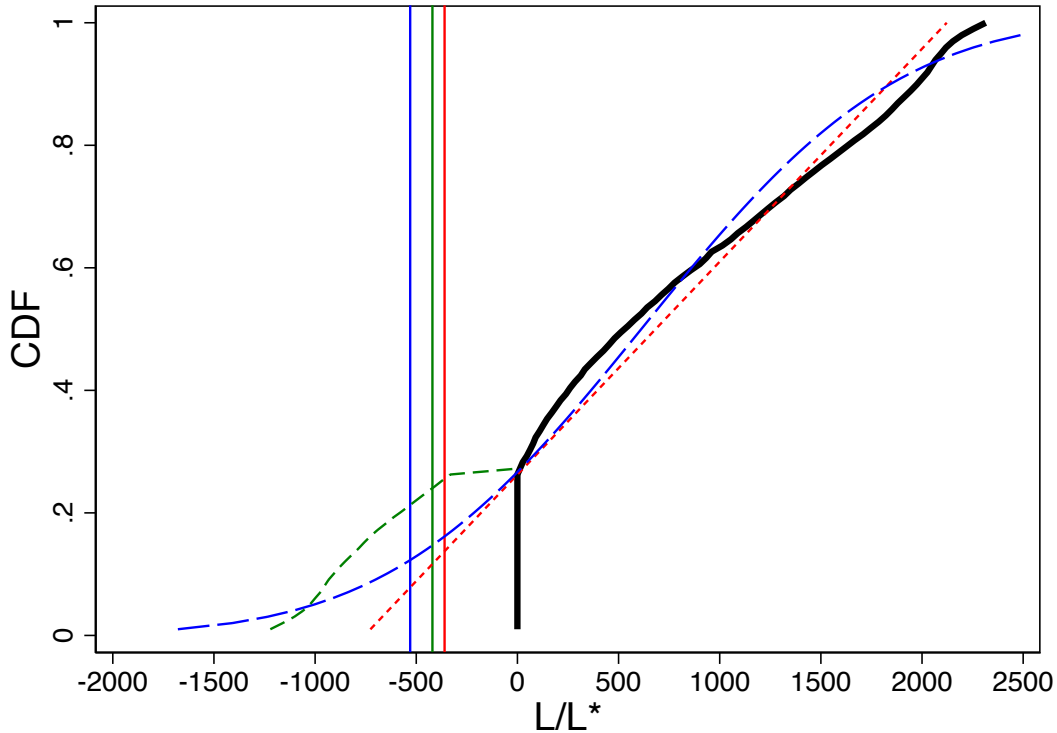
$$S = \beta(L,X) + g(X) + \delta(X)\eta + \varepsilon, \quad \mathbb{E}[\varepsilon|L,X,\eta] = 0$$

$$L^* = h(X) + \eta, \qquad \eta \sim \text{Normal}, \text{Uniform}, \text{ or Symmetric}$$

$$\mathbb{E}[S|L,X] = \beta(L,X) + m(X) + \delta(X)\underbrace{[L + \mathbb{E}[L^*|L=0,X]1(L=0)]}_{\text{new regressor}}$$

$\mathbb{E}[L^*|L=0,X]$

▸ Test 1 ▸ Test 2 ▸ Test 3



$$S = \beta(L,X) + g(X) + \delta(X)\eta + \varepsilon, \quad \mathbb{E}[\varepsilon|L,X,\eta] = 0$$

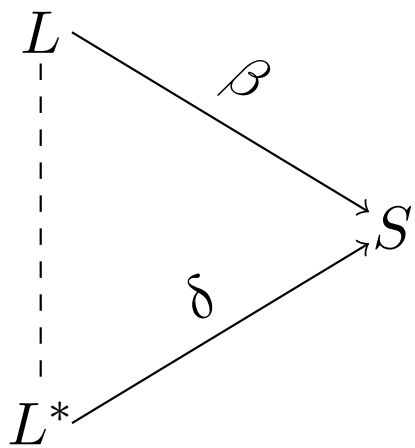
$$L^* = h(X) + \eta, \qquad \eta \sim \text{Normal, Uniform, or Symmetric}$$

$$\mathbb{E}[S|L,X] = \beta(L,X) + m(X) + \delta(X)\underbrace{[L + \mathbb{E}[L^*|L=0,X]1(L=0)]}_{\text{new regressor}}$$

Relation to RDD

Back

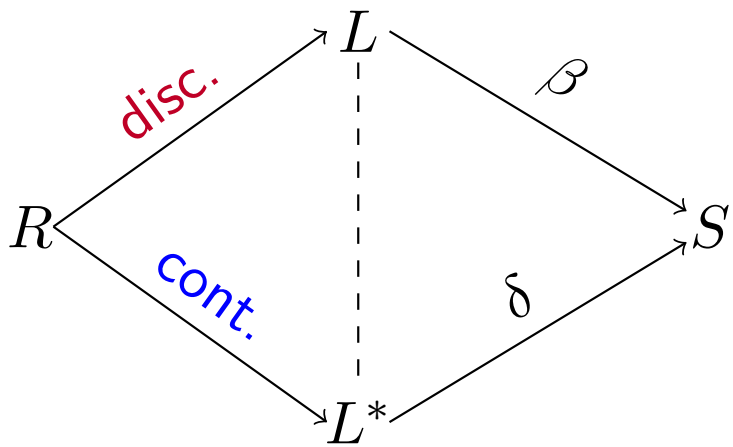
$$S = \alpha + \beta L + \delta L^* + \epsilon$$



Relation to RDD

Back

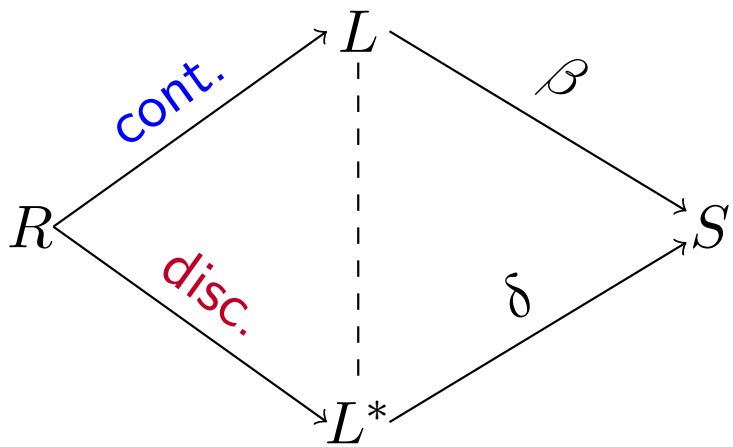
$$S = \alpha + \beta L + \delta L^* + \epsilon$$



Relation to RDD

Back

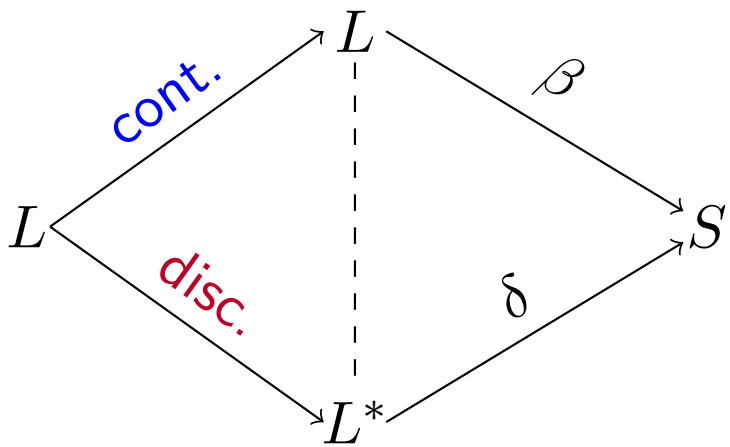
$$S = \alpha + \beta L + \delta L^* + \epsilon$$



Relation to RDD

Back

$$S = \alpha + \beta L + \delta L^* + \epsilon$$
$$L = \max\{0, L^*\}$$

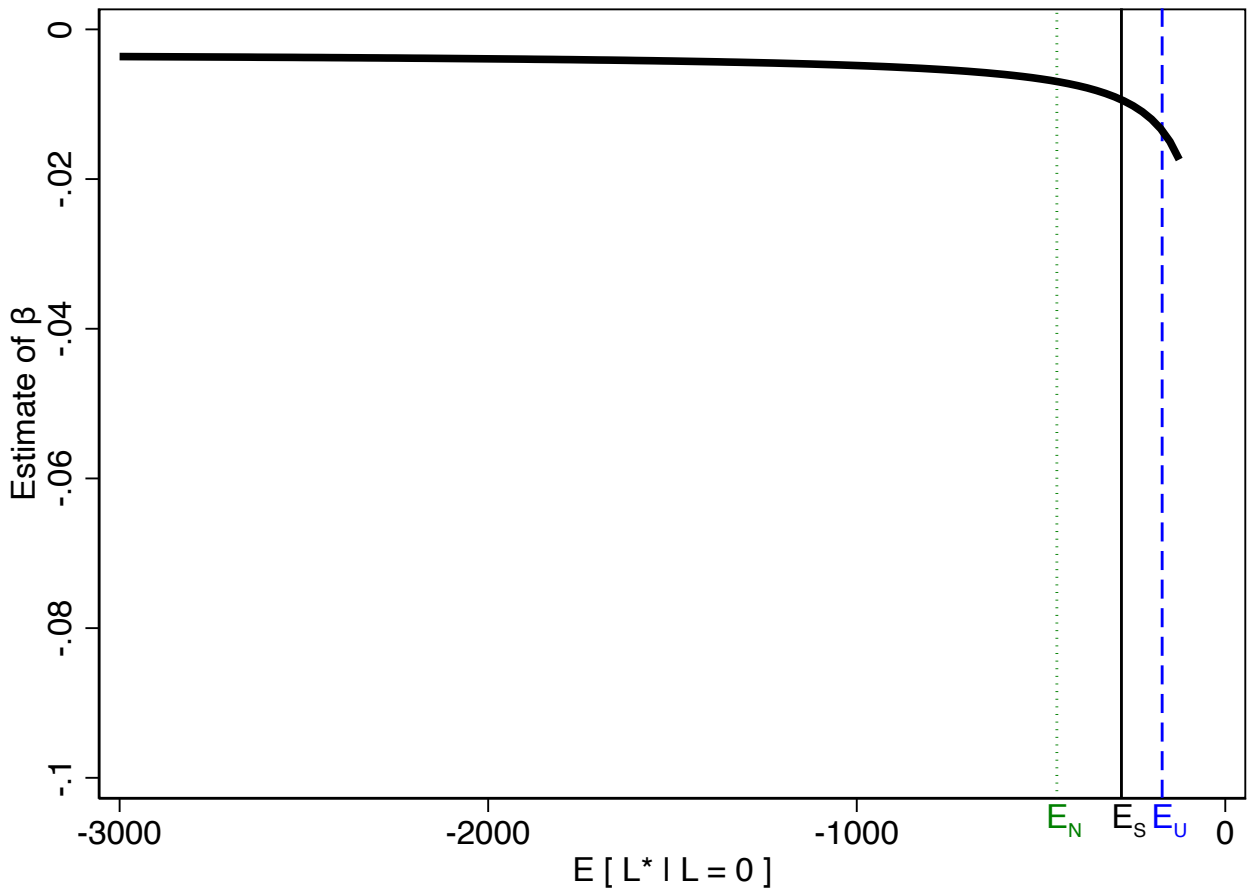


Treatment: $\lim_{l \to 0^+} \mathbb{E}[\beta L | L = l] - \mathbb{E}[\beta L | L = 0] = 0$

Endogeneity: $\lim_{l \to 0^+} \mathbb{E}[\delta L^* | L = l] - \mathbb{E}[\delta L^* | L = 0] = -\delta \mathbb{E}[L^* | L = 0]$

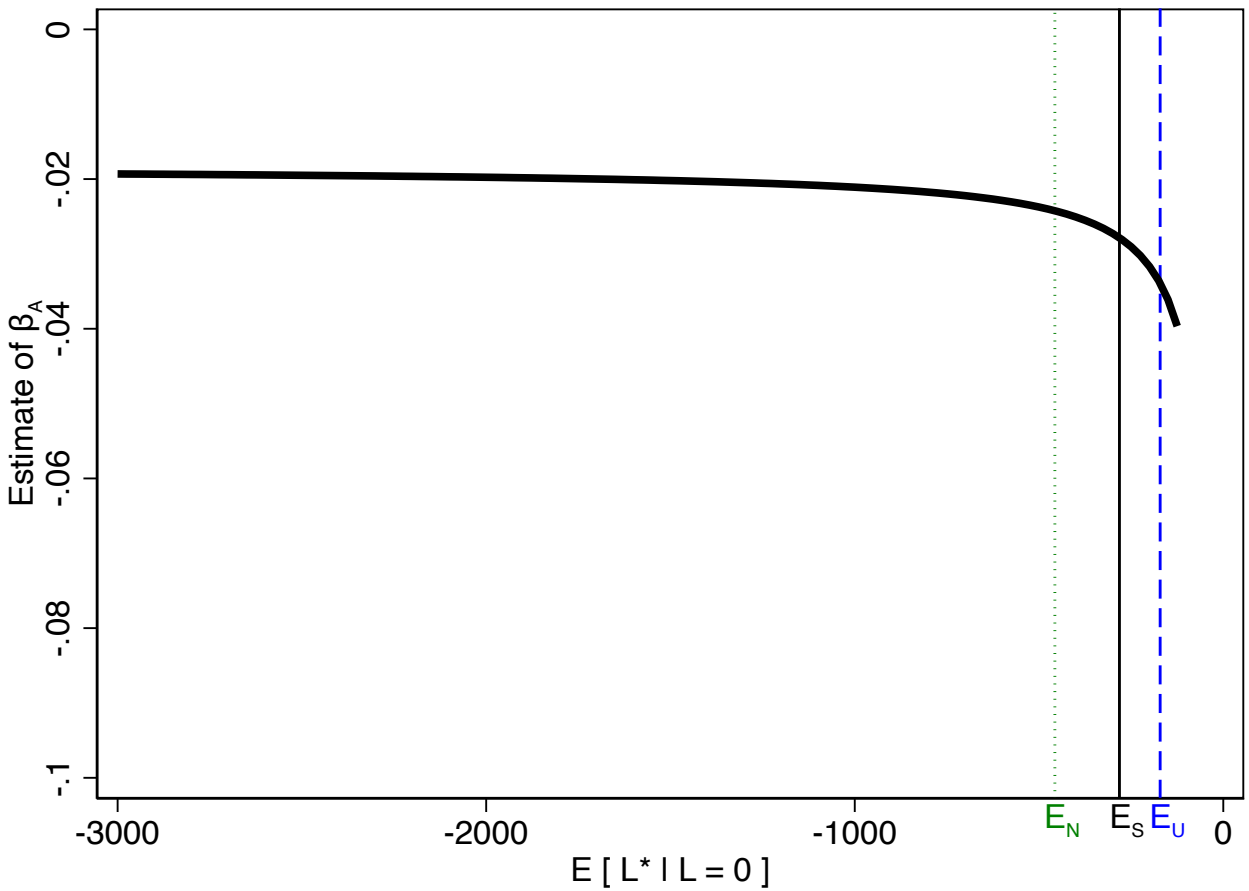
Robustness to Distributional Assumption

Back



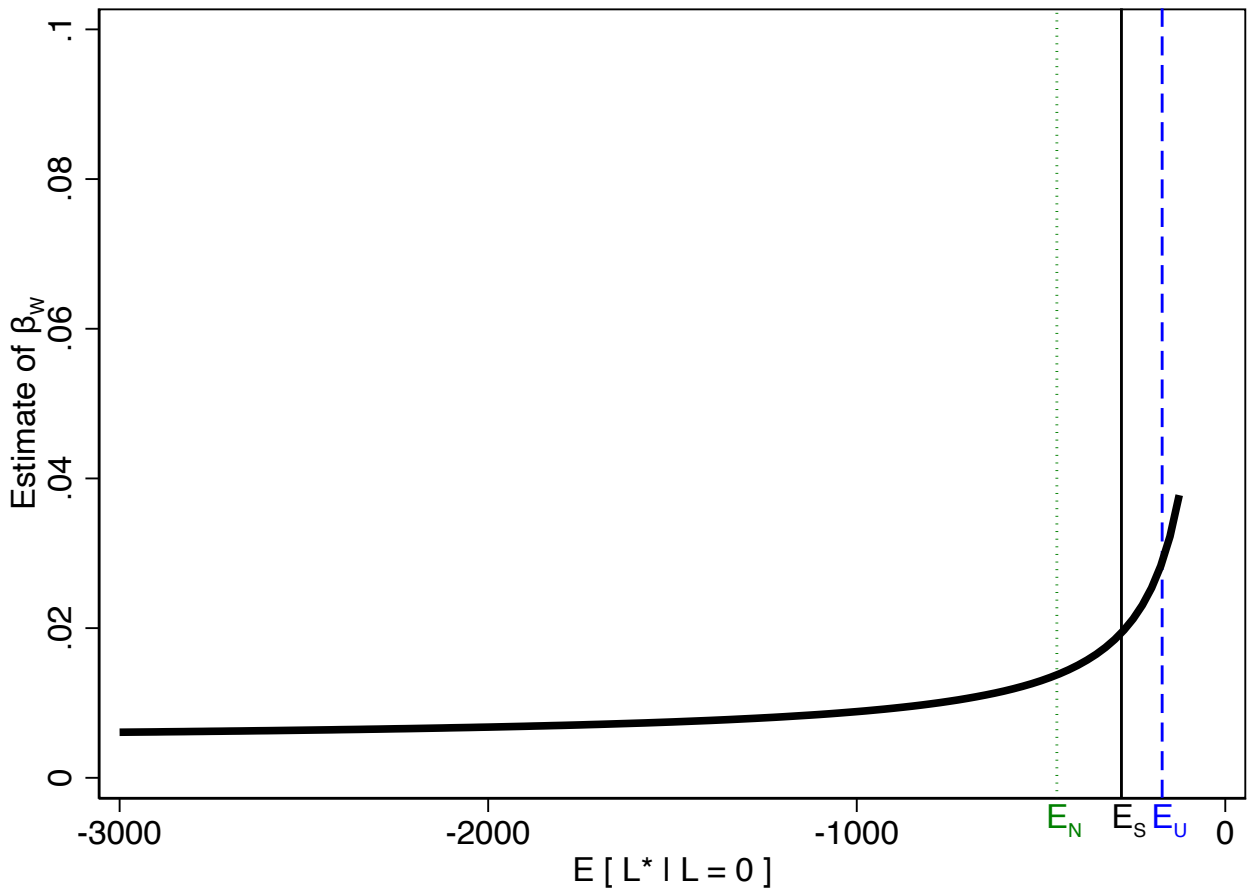
Robustness to Distributional Assumption

Back



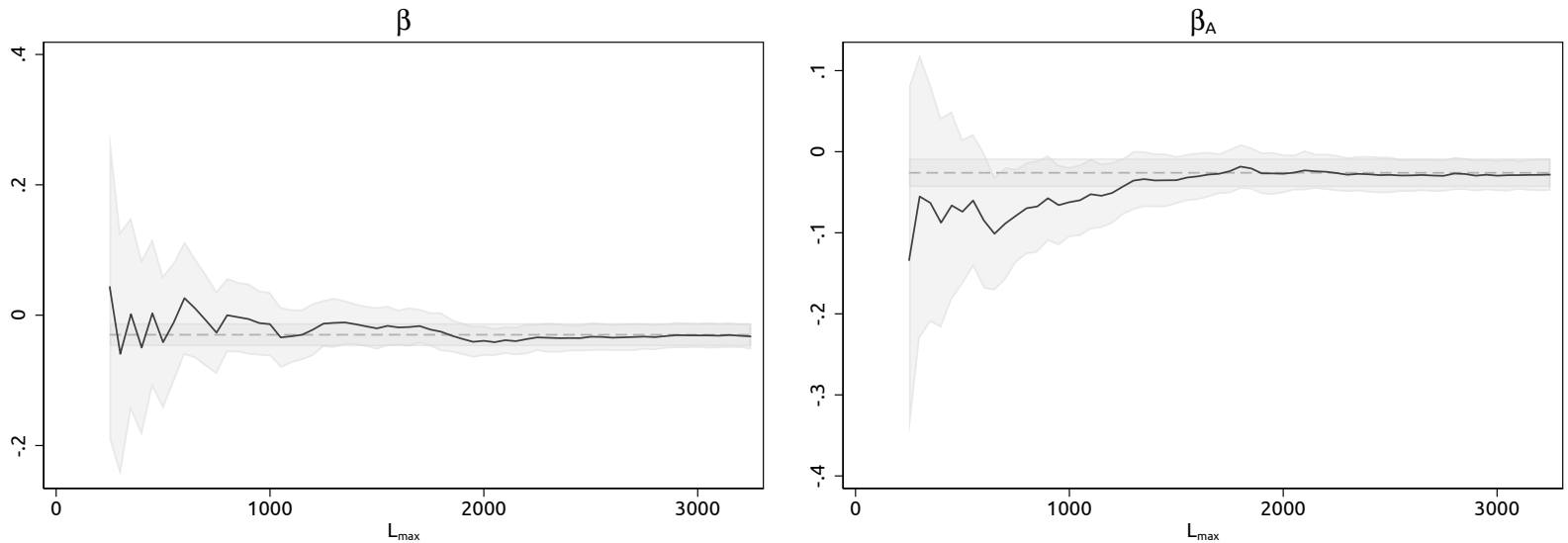
Robustness to Distributional Assumption

Back



Testing Assumption $\mathbb{E}[\varepsilon|L, X, \eta] = 0$

Back



Estimation Details

- It is a good idea to discretize X first, then use cluster indicators to approximate $m(X)$, $\delta(X)$ and $\mathbb{E}[L^*|L^* \leq 0, X]$.
- We use hierarchical cluster with 50 clusters in all reported results.
- Intuition: as the number of clusters increase, these approximations improve.
- See Caetano, Caetano and Nielsen (2021) for details.

Changing the number of clusters

Back