Estimating marginal returns to medical care: Evidence from at-risk newborns

Douglas Almond, Columbia
Joseph Doyle, MIT
Amanda Kowalski, Yale
Heidi Williams, Harvard

NBER Health Care meeting
24 July 2009

Motivation

- Medical expenditures are high, increasing
- Do the benefits of additional expenditures exceed their costs?
  - Estimates of marginal returns needed to inform policy
- Empirical challenge: sicker patients usually receive more inputs
- This paper:
  - Idea: Compare inputs, outcomes across diagnostic thresholds
  - Our focus: 1500 gram "very low birth weight" threshold
    - Just under 3 pounds, 5 ounces

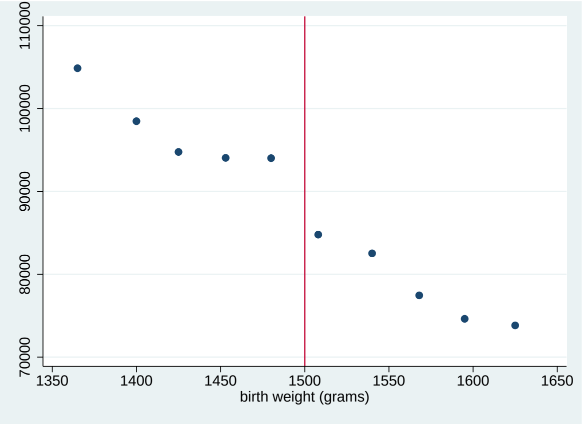
Why study at-risk newborns?

- High value of even small mortality reductions
- Recent high-cost expansions in treatment
  - Our time period: 1980s-2000s
- Little evidence on returns to incremental expenditures
  - Average returns appear high (Cutler & Meara 2000)
  - Dearth of randomized trials
- Substantively large portion of health care system
  - Child birth common and expensive
  - Data on census of births available
    - Large sample size to test for effects on mortality

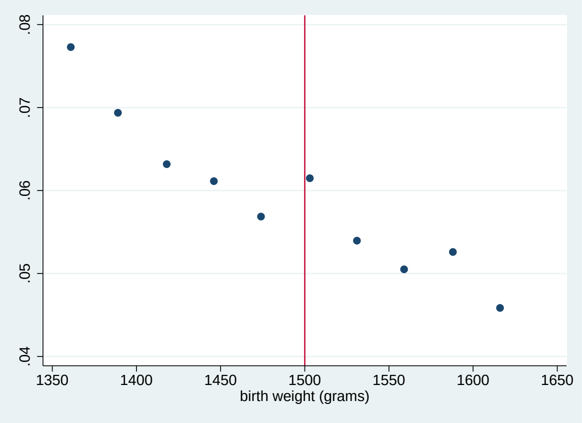
Why study birth weight thresholds?

- Do not represent "real" breaks in underlying health risk
  - *"...designation of very low birth weight infants as those weighing 1,500 grams or less reflected convention rather than biologic criteria"* (Institute of Medicine 1985)
- Birth weight very near threshold "as good as random"
  - Pre-delivery birth weight estimates imprecise
  - Empirically, no evidence of small-scale manipulation
- *A priori* plausible to affect treatment
  - Familiar to physicians (*e.g.* Cloherty & Stark)
  - Reflected in diagnosis, billing codes
  - Empirically, appear to affect treatment provision

Preview of results: Hospital charges



Preview of results: One-year mortality



Preview of results: Estimating marginal returns

- Take hospital charges as best available summary measure
- Scale hospital charges by cost-to-charge ratio
- Assume costs fully capture impact of "VLBW" on mortality

- Implied cost of saving statistical life of a newborn with birth weight ~1500 grams is on the order of $550,000 (2006 dollars)

Data: Mortality by birth weight

Mortality by birth weight: "*Nationwide data*"

- NCHS linked birth & death certificate data
- Census of US births: 1983-1991, 1995-2002
- $\sim$66 million total births
- $\sim$200,000 births within 3oz of 1500 grams
- Rich set of covariates (*e.g.* gestational age, mother's age)

Data: Treatment by birth weight

Treatment by birth weight: "*Five-state sample*"

- CA OSHPD longitudinal database
  - Census of CA birth hospital discharges: 1991-2002
  - Linked to death certificates
- HCUP State Inpatient Databases
  - AZ, NJ, NY, & MD
  - Census of birth hospital discharges: various years, 1991-2006
  - *Not* linked to death certificates; mortality from nationwide data
- Combined five-state sample
  - ∼10.5 million total births
  - ∼30,000 births within 3oz of 1500 grams
  - Charges, length of stay, procedure and diagnosis codes

Additional datasets

- CMS: Hospital-year cost-to-charge ratios
- AHA annual surveys: NICU availability
- Phibbs *et al.* (2007): CA NICU availability by quality level

Summary statistics

Summary statistics in 3oz above 1500 grams ("control group")

- Nationwide data, 1-year mortality: 5.5%
- Five-state sample, hospital charges: $81,500
- Five-state sample, length of stay: 25 days

Empirical framework

Standard methods for regression discontinuity analysis
(*e.g.* Imbens & Lemieux (2008), Lee & Lemieux (2009))

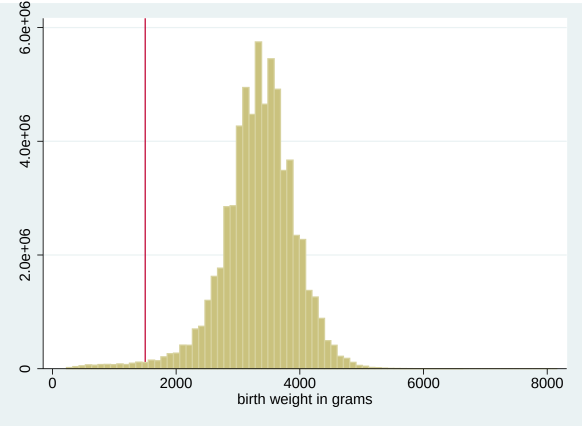
- Pilot bandwidth: 3oz (85 grams)
  - Results qualitatively similar across a wide range of bandwidths

Empirical framework (2)

- Local linear regressions
  - Triangle kernal: weight decays with distance from 1500 grams
  - Asymptotic standard errors
    - Cheng *et al.* (1997), Porter (2003)

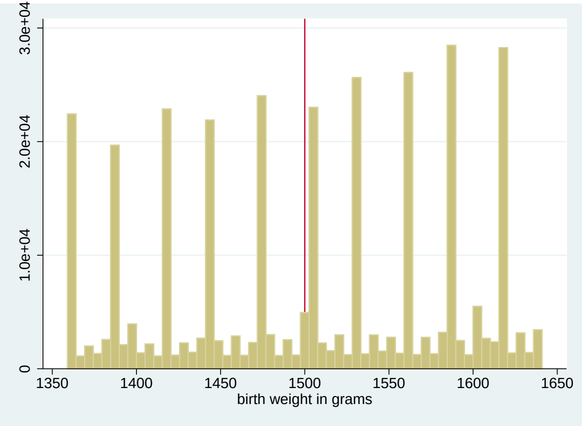
- OLS
  - For infant i weighing grams g in state s in year t:
    $Y_i = \alpha_0 + \alpha_1(VLBW)_i$
    $+\alpha_2(VLBW)_i * (g_i - 1500) + \alpha_3(1 - VLBW)_i * (g_i - 1500)$
    $+\alpha_4(t_i) + \alpha_5(s_i) + \alpha_6(X_i') + \epsilon_i$
  - Robust and Card-Lee (2008) standard errors

Frequency of births: Full sample



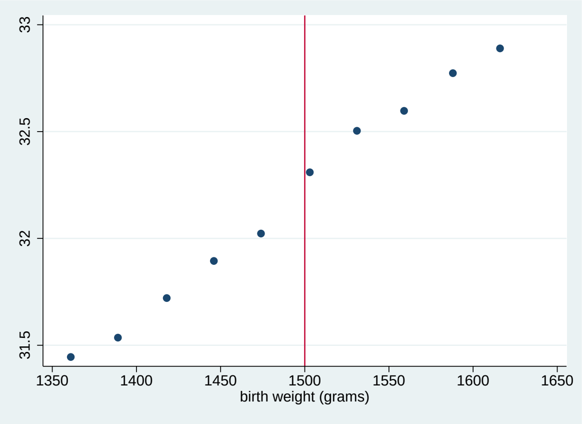
Notes: Nationwide data.

Frequency of births: 1350-1650 grams



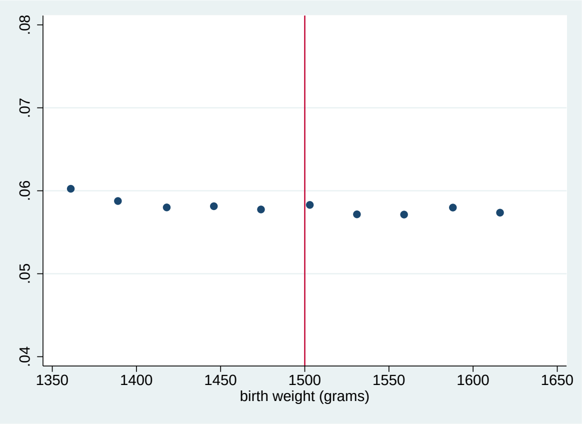
Notes: Nationwide data.

Gestational age



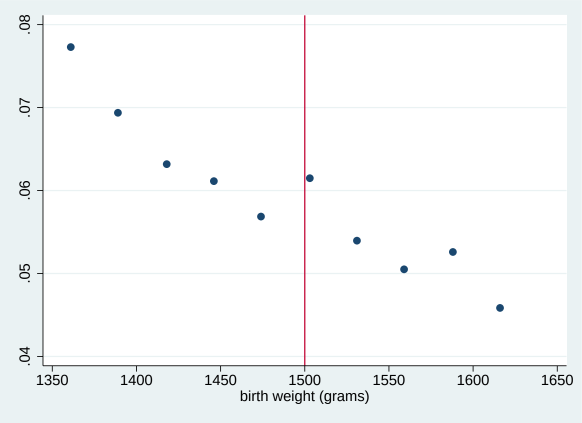
Notes: Nationwide data.

Predicted one-year mortality



Notes: Nationwide data.

One-year mortality: Nationwide data

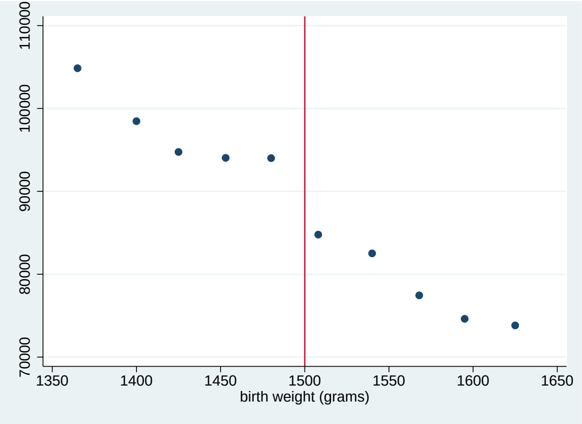


One-year mortality: Nationwide data

Dependent variable:	one-year mortality			
Model:	local linear	OLS	OLS	OLS
Birth weight < 1500g	-0.0121	-0.0095	-0.0076	-0.0072
	(0.0023)**	(0.0022)**	(0.0022)**	(0.0022)**
		[0.0048]*	[0.0040]	[0.0040]
Year controls		No	Yes	Yes
Main controls		No	No	Yes
Mean of dependent variable above cutoff:	0.0553			
Observations	202071			

Notes: OLS models also include (not shown) linear trends interacted with the treatment indicator. Local linear models report asymptotic standard errors. OLS models report heteroskedastic-robust standard errors in parentheses, and standard errors clustered at the gram level in brackets. *: $p<0.05$; **: $p<0.01$.

Hospital charges: Five-state sample

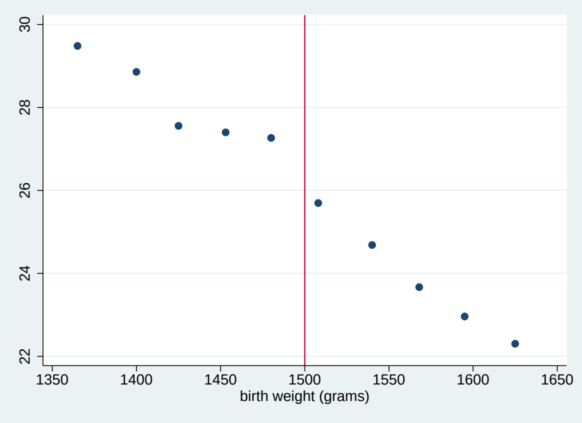


Hospital charges: Five-state sample

Dependent variable:	hospital charges			
Model:	local linear	OLS	OLS	OLS
Birth weight < 1500g	9450	9,022	8,205	9,065
	(2710)**	(2,448)**	(2,416)**	(2,297)**
		[3,538]*	[3,174]*	[5,094]
Year controls		No	Yes	Yes
Main controls		No	No	Yes
Mean of dependent variable above cutoff:	81566			
Observations	28928			

Notes: OLS models also include (not shown) linear trends interacted with the treatment indicator. Local linear models report asymptotic standard errors. OLS models report heteroskedastic-robust standard errors in parentheses, and standard errors clustered at the gram level in brackets. *: $p<0.05$; **: $p<0.01$. Charges are in 2006 dollars.

Length of stay: Five-state sample



Length of stay: Five-state sample

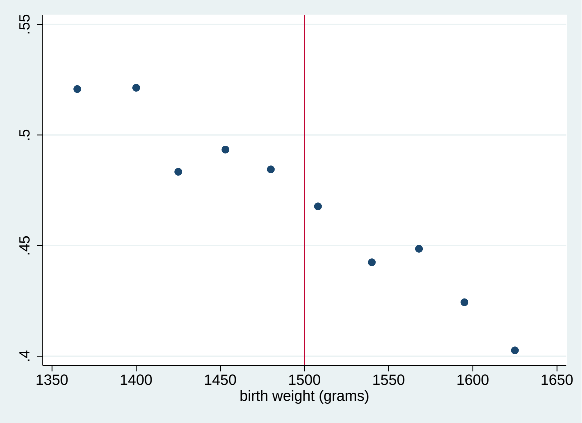
Dependent variable:	length of stay			
Model:	local linear	OLS	OLS	OLS
Birth weight < 1500g	1.97	1.7768	1.7600	1.4635
	(0.451)**	(0.4165)**	(0.4166)**	(0.4107)**
		[1.0024]	[0.9775]	[0.7928]
Year controls		No	Yes	Yes
Main controls		No	No	Yes
Mean of dependent variable above cutoff:	24.68			
Observations	30935			

Notes: OLS models also include (not shown) linear trends interacted with the treatment indicator. Local linear models report asymptotic standard errors. OLS models report heteroskedastic-robust standard errors in parentheses, and standard errors clustered at the gram level in brackets. *: $p<0.05$; **: $p<0.01$.

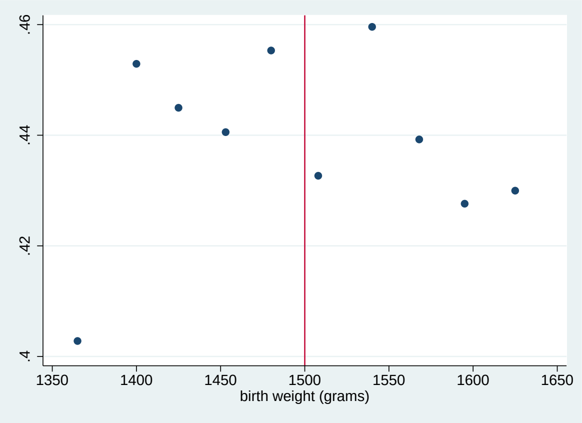
Mechanisms: Differences in types of care

- Procedure codes, focus on common perinatal procedures
  - *E.g.* Ventilation, NICU admission
- Difficult to find differences in five-state sample
  - Likely lack statistical power to detect differences
- Some weak evidence for two relatively common procedures:
  - Diagnostic ultrasounds (consistent with *e.g.* Cloherty & Stark)
  - Operations on the heart

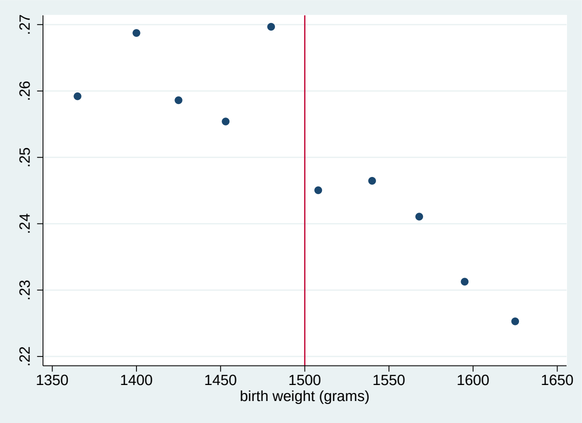
Mechanisms: Any ventilation, five-state sample



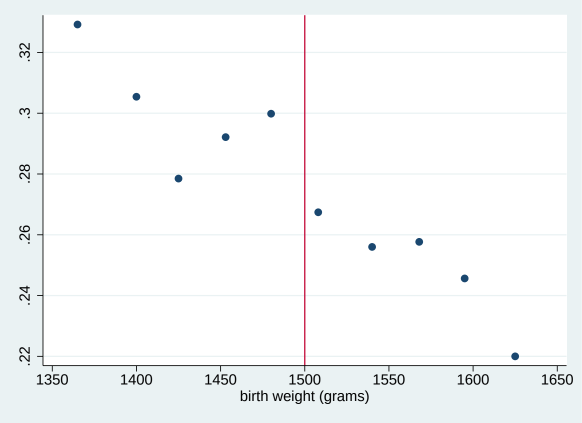
Mechanisms: NICU > 24 hours, five-state sample



Mechanisms: Diagnostic ultrasound, five-state sample



Mechanisms: Operations on the heart, five-state sample



Robustness & specification tests

- Bandwidth sensitivity
  - Results qualitatively similar for a wide range of bandwidths
- Polynomial order sensitivity
  - Results qualitatively similar for higher-order polynomials
- Alternative first stage outcomes
  - log(charges), median charges, including transfers
- Causes of death
  - No statistically significant change in external deaths
  - Largest effects for perinatal conditions (*e.g.* jaundice)

Alternative thresholds

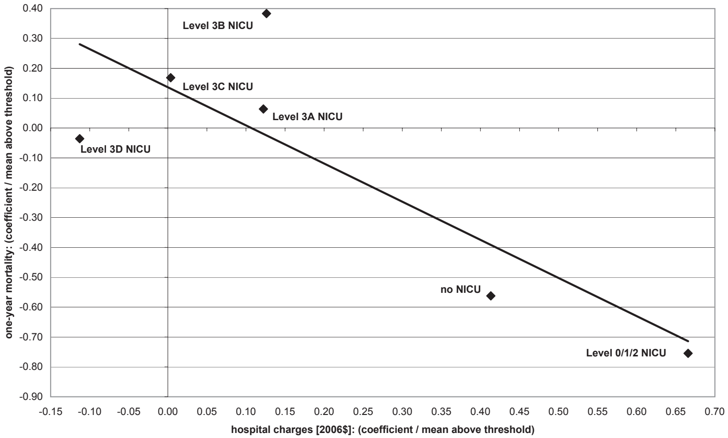
Alternative thresholds: Robustness & alternative estimates

- Unexpected jumps: no other convincing differences
- No convincing differences at 2500 grams ("low birth weight")
- No convincing differences at 37 weeks ("premature")
- Some evidence of small-for-gestational-age effects

Results: Summary

- Preliminaries:
  - Birth frequencies: No evidence of heaping
  - Covariates: Generally smooth across the cutoff
- Reduced form: Mortality
  - 1 percentage point lower one-year mortality (mean = 5.5%)
- First stage: Treatment
  - $9,000 increase in hospital charges (mean = $81,500)
  - 1.5 additional days length of stay (mean = 25 days)
- Mechanisms: Weak evidence
- Robustness checks

Heterogeneity across hospitals, by NICU quality



Notes: First stage (2006 charges) and reduced form (one-year mortality) coefficients, by NICU level in CA data.

Comparison to time-series estimates of returns

Cutler-Meara style calculation

- Within-birth weight changes in cost, mortality over time
- "Long difference" in CA data, 1991-2002
- Implies cost per newborn life of ~$1 million

Two-sample estimates of marginal returns

- Assume costs fully capture impact of "VLBW" on mortality
- Inoue & Solon (2005) for asymptotic confidence intervals

- Five-state sample: ~\$615,270 per newborn life
  - = \$4,553/0.0074
  - 95% confidence interval: \$30,000 to \$1.20 million
- All available data: ~\$527,083 per newborn life
  - = \$3,795/0.0072
  - 95% confidence interval: \$30,000 to \$1.05 million

Some cost-effectiveness benchmarks

- Disability-adjusted value of newborn life: $2.7 million
  - Cutler & Meara (2000)
- Value of (non-disabled) newborn life: $3-$7 million
  - Cutler (2004)

Conclusions

- Medical inputs may be discontinuous across plausibly smooth measures of health risk
- Estimates relevant to “marginally untreated” sub-population

- Here: $\sim$\$550,000 per statistical life saved (2006 dollars)
- Suggests high returns to medical care for newborns near 1500g