

# **Identification and Sensitivity Analysis in Randomized Experiments with Partial Compliance**

Keisuke Hirano

Department of Economics, University of Arizona

hirano@u.arizona.edu

## Overview

- Randomized experiment, but subjects assigned to treatment group may not comply, or comply only partially.
- Review basic results about identification of treatment effects with “all-or-nothing” compliance.
- Simple extensions to bounds and sensitivity analysis when compliance can be partial.
- Illustration using data from Ialongo et al (1999) and Jo (2002).

We follow the general framework of Angrist, Imbens, and Rubin (1996).

No measure of “compliance” for control group, in contrast to Efron and Feldman (1991).

Sensitivity Analysis: Vansteelandt and Goetghebeur (2003, 2005), Robins and Rotnitzky (2004).

Many other papers on noncompliance, including: Robins (1986), Robins and Tsiatis (1991), Robins and Greenland (1994), and others mentioned below.

## Basic Setup without Covariates:

$Z_i$  : randomized assignment ( $= 0, 1$ )

$D_i(z)$  : potential treatment

Assume  $D_i(z) = 0$ .

Realized treatment:  $D_i = D_i(Z_i)$  (always measured)

$Y_i(d)$  : potential outcome/dose response.

Exclusion restriction: no direct effect of assignment—see Frangakis (1999), Hirano, Imbens, Rubin, Zhou (2000), Jo (2002).

Average dose-response:  $E[Y_i(d)]$ .

Realized outcome:  $Y_i = Y_i(D_i)$ .

(drop  $i$  subscript below, but keep in mind we are allowing these functions to vary among individuals)

## Binary Treatment (All-or-nothing compliance)

$$D(0) = 0,$$

$$D(1) = 0 \text{ or } 1.$$

We can define two “compliance types”:

- Never-takers:  $D(1) = 0$ .
- Full Compliers:  $D(1) = 1$ .

Let  $C = n, f$  denote never-taker or full complier.

## Identification:

We can estimate  $E[Y|Z, D]$ , and  $Pr[D = 1|Z = 1]$ .

Want to relate this to causal quantities, in particular subgroup dose-response functions.

By random assignment of  $Z$ :

$$E[Y(1)|C = f] = E[Y|Z = 1, D = 1],$$

$$E[Y(0)|C = n] = E[Y|Z = 1, D = 0],$$

and

$$\pi_f := Pr(C = f) = Pr[D = 1|Z = 1].$$

Would like to estimate  $E[Y(0)|C = f]$  and then the average treatment effect among full compliers:

$$\begin{aligned}ATE_f &\equiv E[Y(1) - Y(0)|C = f] \\ &= E[Y(1)|C = f] - E[Y(0)|C = f].\end{aligned}$$

(Cannot hope to estimate  $E[Y(1)|C = n]$ , so cannot point-identify treatment effect for never-takers.)



Note that  $Z = 0$  outcomes are a mixture between  $C = f$  and  $C = n$  groups:

$$\begin{aligned} E[Y|Z = 0] &= E[Y(0)] \\ &= \pi_f E[Y(0)|C = f] + (1 - \pi_f) E[Y(0)|C = n]. \end{aligned}$$

We can estimate  $E[Y|Z = 0]$ , and we also can estimate  $\pi_f$  and  $E[Y(0)|C = n]$ .

Thus  $E[Y(0)|C = f]$  is identified:

$$E[Y(0)|C = f] = \left\{ E[Y|Z = 0] - (1 - \pi_f) E[Y(0)|C = n] \right\} / \pi_f.$$

Can estimate  $E[Y(0)|C = f]$  by replacing expectations and probabilities with sample versions, or more sophisticated methods.

In practice may want to incorporate covariates and further modeling/smoothing assumptions.

“Nonparametric” identification: data are fundamentally informative about  $ATE_f$ ; we are not relying on auxiliary functional form assumptions.

## Empirical Example

Johns Hopkins Public School Preventive Intervention Study (Ialongo, Werthamer, Kellam, Brown, Wang, and Lin, 1999) (Jo, 2002)

Randomized evaluation of program to improve first-grade child achievement and behavior.

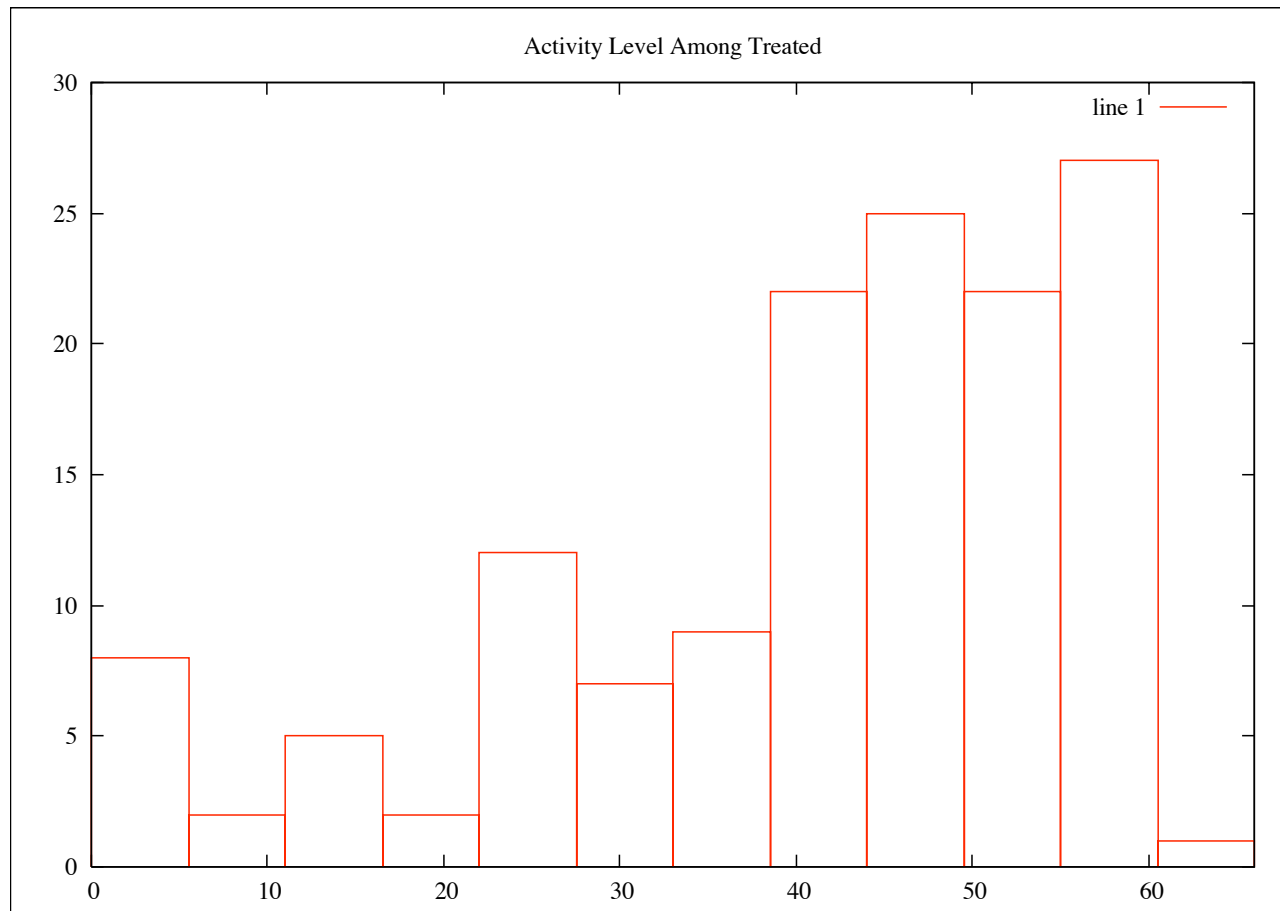
Intervention: parent asked to perform 66 take-home activities.

Sample size: 284.

Outcome: **SHY18** - composite of TOCA-R scores related to shy behavior 18 months after start of program - a number between 1 and 6.

Note: there was also a missing data problem - handled here by listwise deletion, but other methods may be preferable.

## Compliance among Z=1 Group



Intention-to-Treat Effects:

$Z = 0$	$Z = 1$	Difference in Means
2.46	2.15	-0.31
(0.09)	(0.08)	(0.12)

To start, define “full compliance” as:

$$D = 1(\text{Activities} \geq 45).$$

Sample means:

	$D = 0$	$D = 1$
$Z = 0$	2.46 (.09)	–
$Z = 1$	2.03 (.12)	2.27 (.12)

Linear IV estimates:

	(1)	(2)
Const	2.46 (.09)	1.99 (.16)
D	-0.61 (.25)	-.67 (.24)
Shy0		0.22 (.06)

(Coefficient on  $D$  consistently estimates  $ATE_f$ .)

## Alternative definition of full compliance

$$D = 1(\text{Activities} \geq 54).$$

Sample means:

	$D = 0$	$D = 1$
$Z = 0$	2.46 (.09)	—
$Z = 1$	2.18 (.10)	2.07 (.14)

Linear IV estimates:

Const	D
2.46 (.09)	-1.24 (.51)



### 3-valued Treatment

$$D(0) = 0,$$

$$D(1) = 0, \frac{1}{2}, 1.$$

3 compliance types:

$D(1) = 0$ : never-takers ( $C = n$ ),

$D(1) = \frac{1}{2}$ : partial compliers ( $C = p$ ),

$D(1) = 1$ : full compliers ( $C = f$ ).

Potential outcomes:  $Y(0), Y(1/2), Y(1)$ .

**Identified quantities:**

$$\pi_n = Pr(C = n) = E[1(D = 0)|Z = 1],$$

$$\pi_p = Pr(C = p) = E[1(D = \frac{1}{2})|Z = 1],$$

$$\pi_f = Pr(C = f) = E[1(D = 1)|Z = 1].$$

Also identified:

$$E[Y(1)|C = f] = E[Y|Z = 1, D = 1],$$

$$E[Y(\tfrac{1}{2})|C = p] = E[Y|Z = 1, D = \tfrac{1}{2}],$$

$$E[Y(0)|C = n] = E[Y|Z = 1, D = 0].$$

So we might try to estimate

$$E[Y(1) - Y(0)|C = f],$$

or

$$E[Y(\frac{1}{2}) - Y(0)|C = p].$$

Now

$$E[Y|Z = 0] =$$

$$\pi_n E[Y(0)|C = n] + \pi_p E[Y(0)|C = p] + \pi_f E[Y(0)|C = f]$$

As before, this is a mixture across the different compliance types.

Since we can estimate  $E[Y|Z = 0]$ ,  $E[Y(0)|C = n]$ , and compliance probabilities, the data identify a linear relationship between

$$E[Y(0)|C = p] \text{ and } E[Y(0)|C = f].$$

Suppose that  $Y$  is bounded:  $Y \in [\underline{y}, \bar{y}]$ .

Then this restricts the set of possible values for  $E[Y(0)|C = p]$  and  $E[Y(0)|C = f]$ .

## Empirical Example

Define:

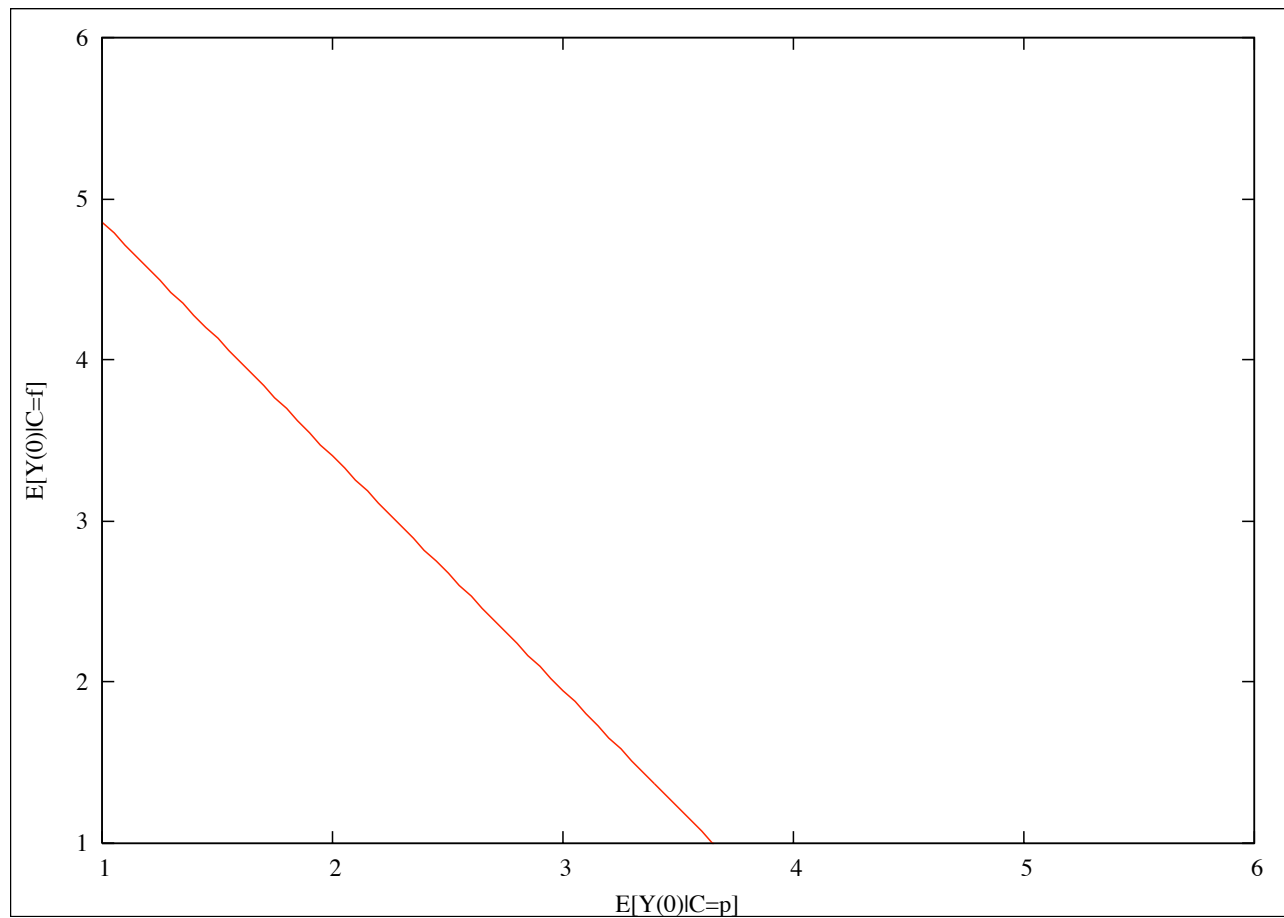
$$\begin{aligned} D &= 0 && \text{if } \text{Activities} \leq 30, \\ D &= \frac{1}{2} && \text{if } 30 < \text{Activities} \leq 50, \\ D &= 1 && \text{if } \text{Activities} > 50. \end{aligned}$$

Sample averages:

$$\begin{aligned}\hat{E}[Y(0)|C = n] &= \hat{E}[Y|Z = 1, D = 0] = 2.1029, \\ \hat{E}\left[Y\left(\frac{1}{2}\right) | C = p\right] &= \hat{E}\left[Y | Z = 1, D = \frac{1}{2}\right] = 2.2097, \\ \hat{E}[Y(1)|C = f] &= \hat{E}[Y|Z = 1, D = 1] = 2.1136. \\ \hat{E}[Y|Z = 0] &= 2.4595.\end{aligned}$$

Use these along with constraint that outcome is between 1 and 6 to obtain implied values of  $E[Y(0)|C = p]$  and  $E[Y(0)|C = f]$ :

**Possible values of  $E[Y(0)|C = p]$  and  $E[Y(0)|C = f]$**





Some possible ways to narrow bounds:

- Assume that  $E[Y(0)|C = p]$  and  $E[Y(0)|C = f]$  do not differ by more than some  $\epsilon$ .
- Incorporate covariates + functional form restrictions.

## Multivalued/Continuous Treatment

$$Z = 0, 1$$

$$D(0) = 0$$

$$D(1) \in \mathcal{T}.$$

Example:  $\mathcal{T} = [0, 1]$  when treatment is a fraction of nominal dosage.

In Ialongo et al data:  $\mathcal{T} = \{0, 1, 2, \dots, 66\}$  activities.

Compliance type:  $C \in \mathcal{T}$  denotes the treatment level the individual would take, IF they were assigned to the treatment arm.

Let  $F_C$  denote the distribution of compliance type  $C$ , with support a subset of  $\mathcal{T}$ .

## Identification and Sensitivity Analysis

$$F_C(c) = P(D \leq c | Z = 1),$$

$$E[Y | Z = 1, D = d] = E[Y(d) | C = d],$$

$$E[Y | Z = 0] = \int_T E[Y(0) | C = c] dF_C(c).$$

Suppose

$$E[Y(0) | C = c] = h(c, \alpha),$$

where  $h$  is a known function and  $\alpha$  is a parameter. Then

$$E[Y | Z = 0] = \int h(c, \alpha) dF_C(c).$$

## Estimation of bounds

Replace  $E[Y|Z = 0]$  and  $F_C$  by sample analogs.

$$\widehat{E[Y|Z = 0]} = \frac{\sum_i Y_i 1(Z_i = 0)}{\sum_i 1(Z_i = 0)},$$

$$\int h(c, \alpha) d\hat{F}_C(c) = \frac{\sum_i h(D_i, \alpha) 1(Z_i = 1)}{\sum_i 1(Z_i = 1)}.$$

Then choose  $\hat{\alpha}$  to solve

$$\widehat{E[Y|Z = 0]} = \int h(c, \hat{\alpha}) d\hat{F}_C(c),$$

s.t. constraint that  $h(c, \hat{\alpha}) \in [\underline{y}, \bar{y}]$ .

If  $\dim(\alpha) > 1$ , then solution will typically be set-valued.

## Large-sample theory

No smoothing used, so this set-valued estimator should be relatively well-behaved.

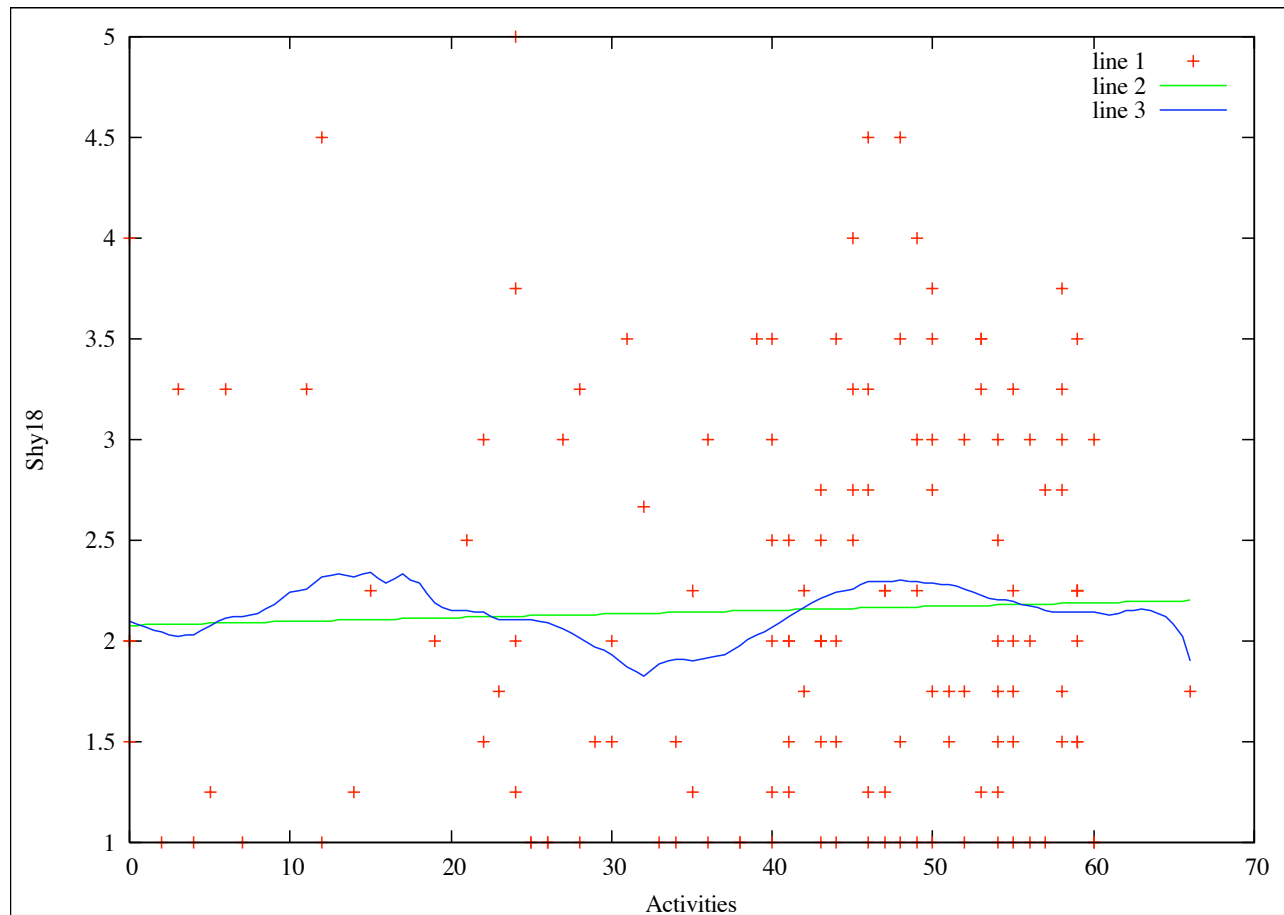
Theory for set-valued parameter estimates under rapid construction:

Manski and Tamer: consistency wrt Hausdorff metric

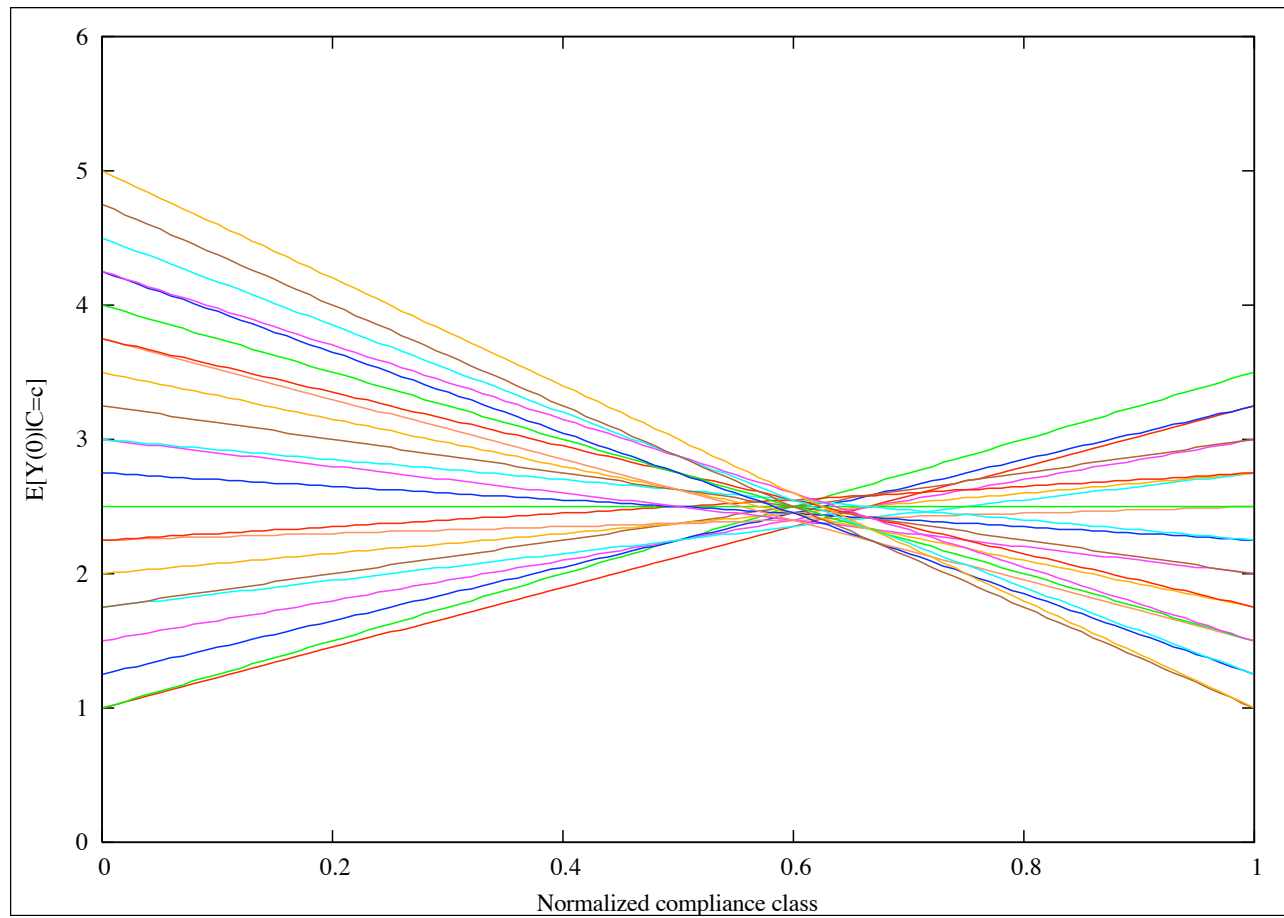
Imbens and Manski: coverage properties

Chernozhukov, Hong, Tamer: large-sample distribution theory for inference  
etc

## Dose-response for Z=1: Linear and Kernel Regression



## Sensitivity Analysis: $h(c, \alpha)$ Linear





## **Extensions: Incorporating covariates**

Under functional form/smoothing/prior restrictions, can gain identification, but this can be sensitive to assumptions:

Frangakis (1999), Hirano, Imbens, Rubin, Zhou (2000), Vansteelandt and Goetghebeur (2003, 2005), Robins and Rotnitzky (2004).

Here, could use covariates + functional form restrictions to further narrow bounds.

## **Extensions: Missing Outcome Data**

Frangakis and Rubin (1999), Yau and Little (2001), Barnard et al (2003), Mealli and Rubin (2003), O'Malley and Normand (2005).

For this data set, see discussion of Jo (2002).

## **Design of experiments**

Could assign different dosages to different individuals, or also randomize some additional “encouragements”

In some cases this may help identification: recent work on ID with continuous assignment and treatment includes Imbens and Newey (2004).

## Conclusion:

- Potential outcomes to define causal quantities (dose response functions).
- Identification analysis: connecting estimable quantities to counterfactual/causal quantities, without relying on strong functional form assumptions.
- Partial noncompliance restricts what can be learned without auxiliary assumptions.
- Next step is to introduce further subject-specific assumptions to narrow bounds and sharpen inference.