Bounds analysis of unemployment durations in Germany

Sokbae (Simon) Lee

IFS and UCL

and

Ralf Wilke

ZEW Mannheim

Introduction

- This project is concerned with estimating effects of the reform of the German unemployment compensation system during the 1980s.
- Natural experiment: for elderly unemployed workers, the maximum entitlement periods for unemployment benefits have increased during the mid 1980s. See Hunt (1995) and Fitzenberger and Wilke (2004).

- Early retirement issue for aged >50: Sample composition depends on treatment. Restrict analysis to aged <50.
- Hunt (1995) found evidence for significant effects of extension of unemployment benefits using German Socioeconomic Panel (survey data). Fitzenberger/Wilke (2004) found that results depend on the definition of unemployment using IAB employment subsample (IABS, register data).

- The goal of this project is to revisit Hunt's findings using IABS 1975-1997.
- In particular we aim to provide robust results in terms of definition of unemployment and to overcome possible selection problems.

Data

- IAB employment subsample is register based data of individuals of about 500,000 individuals in West-Germany.
- This has much larger sample size compared to German Socioeconomic Panel, but unemployment spells are not identified from the data.
- The data contains daily information on employment spells and the spells during which individuals receive transfer payments from the labor offices.
- We expect less measurement error regarding the unemployment spells than in survey data.

Data (continued)

- Proxies for unemployment (Fitzenberger/Wilke, 2004)

- Nonemployment (NE): All periods of nonemployment during which the individual receives at least for one day income transfer payments.

- Unemployment between jobs (UBJ): Periods between employment spells during which an individual continuously receives income transfer payments.

- NE is an upper bound of the unobserved unemployment spell and UBJ is a lower bound.

Data (continued)

• Reform years: 1985-1987

• Pre-reform years: 1981 - 1983 (3 years)

- Unemployment spells starting in 1983 are the latest not affected by the reform.

• Post-reform years: 1987 - 1994 (8 years)

- The post reform system applies to most of the unemployment spells starting in 1987.

• Control group: workers aged 26-41

- Aged <25 excluded because of youth unemployment policy changes.

Data (continued)

• Treatment group: workers aged 44-48

- Aged 42-43 are excluded because the treatment for this group is weak.

- Aged >48 are not considered because of early retirement issue.

• Sample restricted to:

- only periods with unemployment benefits as first income transfer

- only individuals not receiving any unemployment transfer during the past 12 months and who did not get a recall to the former employer after the last unemployment period.

- only individuals with completed apprenticeship or university degree

- business sector agriculture is excluded

Data (continued)

Table 1: Descriptive summary of the sample: pre reform years

	aged 26-41 (control group)	aged 44-48 (treatment group)
number of spells	8,194	1,481
mean/median spell length UBJ	101/18	114/44
mean/median spell length NE	664/267	494/197
censored (NE)	12%	16%
female	44%	32%
married	75%	81%
mean age (in years)	32.1	45.8

Data (continued)

Table 2: Descriptive summary of the sample: post reform years

	aged 26-41 (control group)	aged 44-48 (treatment group)
number of spells	20,135	3,271
mean/median spell length UBJ	99/10	114/6
mean/median spell length NE	463/260	521/307
censored (NE)	21%	28%
female	50%	44%
married	55%	69%
mean age (in years)	32.2	45.9

Econometric Framework

- Our framework is based on bounds analysis (Manski 2003).
- In particular, we present bounds for treatment effects in the context of difference-in-differences.
- Tighter bounds are obtained using some monotonicity and independence assumptions. See Manski and Pepper (2000) and Blundell, Gosling, Ichimura, and Meghir (2004) among others.
- No new ideas; however, our project appears to be a first application of bounds analysis to duration analysis and difference-in-differences.

Interval data on durations

• Assume that we observe interval data on durations, that is we observe Y_1 and Y_2, where $Y_1 \leq Y_2$. It is only known that latent duration Y is between Y_1 and Y_2. For example, if $Y_1 = Y_2$, then observed duration is a point; however, in general, we have $Y_1 < Y_2$, including standard right-censored cases.

• Define $S(y|x) = P(Y > y|X = x)$, $S_1(y|x) = P(Y_1 > y|X = x)$, and $S_2(y|x) = P(Y_2 > y|X = x)$. With the empirical evidence alone, then the identification region for $S(y|x)$ is

$$S_1(y|x) \leq S(y|x) \leq S_2(y|x). \tag{1}$$

Bounding the treatment effect

- In addition to covariates X, suppose that P denotes time periods 0 and 1 (before and after a reform) and T denotes age groups 0 and 1 (control and treatment groups).

- Consider the treatment effect on the survival probability $S(y|x)$. The effect of a reform can be measured by difference-in-differences.

- Note that the effect of a reform can be estimated by a sample analog of

$$\begin{aligned}\Delta(y|x) = & [S(y|T=1,P=1,X=x) - S(y|T=0,P=1,X=x)] \\ & - [S(y|T=1,P=0,X=x) - S(y|T=0,P=0,X=x)]\end{aligned}$$

Bounding the treatment effect (continued)

- Using the empirical evidence alone,

$$S_1(y|t,p,x) \le S(y|t,p,x) \le S_2(y|t,p,x) \tag{2}$$

for $t = 0, 1$ and $p = 0, 1$.

- Hence,

$$\begin{aligned} S_1(y|1,1,x) - S_2(y|0,1,x) &\le S(y|1,1,x) - S(y|0,1,x) \\ &\le S_2(y|1,1,x) - S_1(y|0,1,x) \end{aligned}$$

and

$$\begin{aligned} S_1(y|1,0,x) - S_2(y|0,0,x) &\le S(y|1,0,x) - S(y|0,0,x) \\ &\le S_2(y|1,0,x) - S_1(y|0,0,x). \end{aligned}$$

Bounding the treatment effect (continued)

- This implies that $\Delta(y|x)$ is bounded by an interval with endpoints $[l(y|x), u(y|x)]$

$$\begin{aligned} l(y|x) = \max[-1, \{S_1(y|1,1,x) - S_2(y|0,1,x)\} \\ - \{S_2(y|1,0,x) - S_1(y|0,0,x)\}] \end{aligned}$$

and

$$\begin{aligned} u(y|x) = \min[1, \{S_2(y|1,1,x) - S_1(y|0,1,x)\} \\ - \{S_1(y|1,0,x) - S_2(y|0,0,x)\}]. \end{aligned}$$

- If this interval is shorter than $[-1, 1]$, there is identifying power. In particular, the lower bound is larger than zero or the upper bound is smaller than zero, then that will provide the sign of the effect.

Imposing additional assumptions

- **[Assumption S1]** Suppose that $S(y|0,p,x) \leq S(y|1,p,x)$ for all p and x.
- This means that young workers tend to have shorter durations than old workers while other things being equal.
- Under this additional assumption,

$$\begin{aligned}\max\{0, S_1(y|1,1,x) - S_2(y|0,1,x)\} &\leq S(y|1,1,x) - S(y|0,1,x) \\ &\leq S_2(y|1,1,x) - S_1(y|0,1,x)\end{aligned}$$

and

$$\begin{aligned}\max\{0, S_1(y|1,0,x) - S_2(y|0,0,x)\} &\leq S(y|1,0,x) - S(y|0,0,x) \\ &\leq S_2(y|1,0,x) - S_1(y|0,0,x).\end{aligned}$$

Imposing additional assumptions (continued)

- This implies that $\Delta(y|x)$ is bounded by an interval with endpoints $[\tilde{l}(y|x), \tilde{u}(y|x)]$

$$\begin{aligned}\tilde{l}(y|x) &= \max[-1, \max\{0, S_1(y|1,1,x) - S_2(y|0,1,x)\} \\ &\quad - \{S_2(y|1,0,x) - S_1(y|0,0,x)\}]\end{aligned}$$

and

$$\begin{aligned}\tilde{u}(y|x) &= \min[1, \{S_2(y|1,1,x) - S_1(y|0,1,x)\} \\ &\quad - \max\{0, S_1(y|1,0,x) - S_2(y|0,0,x)\}].\end{aligned}$$

Imposing additional assumptions (continued)

- **[Assumption S2]** Suppose that $S(y|t,p,x,v) \leq S(y|t,p,x,v')$ for $v \leq v'$.
- In other words, the survivor function is monotone with respect to v. For example, let V be the age of a worker in years. Then this says that the survivor function is increasing as a worker gets older conditional on other variables.
- Then

$$\max_{v' \leq v} S_1(y|t,p,x,v') \leq S(y|t,p,x,v) \leq \min_{v' \geq v} S_2(y|t,p,x,v') \tag{3}$$

for $t = 0, 1$ and $p = 0, 1$.

Imposing additional assumptions (continued)

- By the law of iterative expectations,

$$T_1(y|t,p,x) \leq S(y|t,p,x) \leq T_2(y|t,p,x), \tag{4}$$

where

$$T_1(y|t,p,x) = \sum_v \Pr(V = v|t,p,x)[\max_{v' \leq v} S_1(y|t,p,x,v')]$$

and

$$T_2(y|t,p,x) = \sum_v \Pr(V = v|t,p,x)[\min_{v' \geq v} S_2(y|t,p,x,v')].$$

Imposing additional assumptions (continued)

- Combining this with Assumption S1 gives a tighter bound for $\Delta(y|x)$:

$$\begin{aligned}\bar{l}(y|x) = \max[&\tilde{l}(y|x), \max\{0, T_1(y|1,1,x) - T_2(y|0,1,x)\} \\ &- \{T_2(y|1,0,x) - T_1(y|0,0,x)\}]\end{aligned}$$

and

$$\begin{aligned}\bar{u}(y|x) = \min[&\tilde{u}(y|x), \{T_2(y|1,1,x) - T_1(y|0,1,x)\} \\ &- \max\{0, T_1(y|1,0,x) - T_2(y|0,0,x)\}].\end{aligned}$$

Imposing additional assumptions (continued)

- To further shorten the interval, one may assume that $\Delta(y|x) = \Delta(y)$ for all x in some set $\mathcal{X}$. That is, the treatment effect is independent of $x \in \mathcal{X}$.
- Then

$$\Delta(y) \in [\max_{x \in \mathcal{X}} l(y|x), \min_{x \in \mathcal{X}} u(y|x)]$$

for some upper and lower bounds $l(y|x)$ and $u(y|x)$.

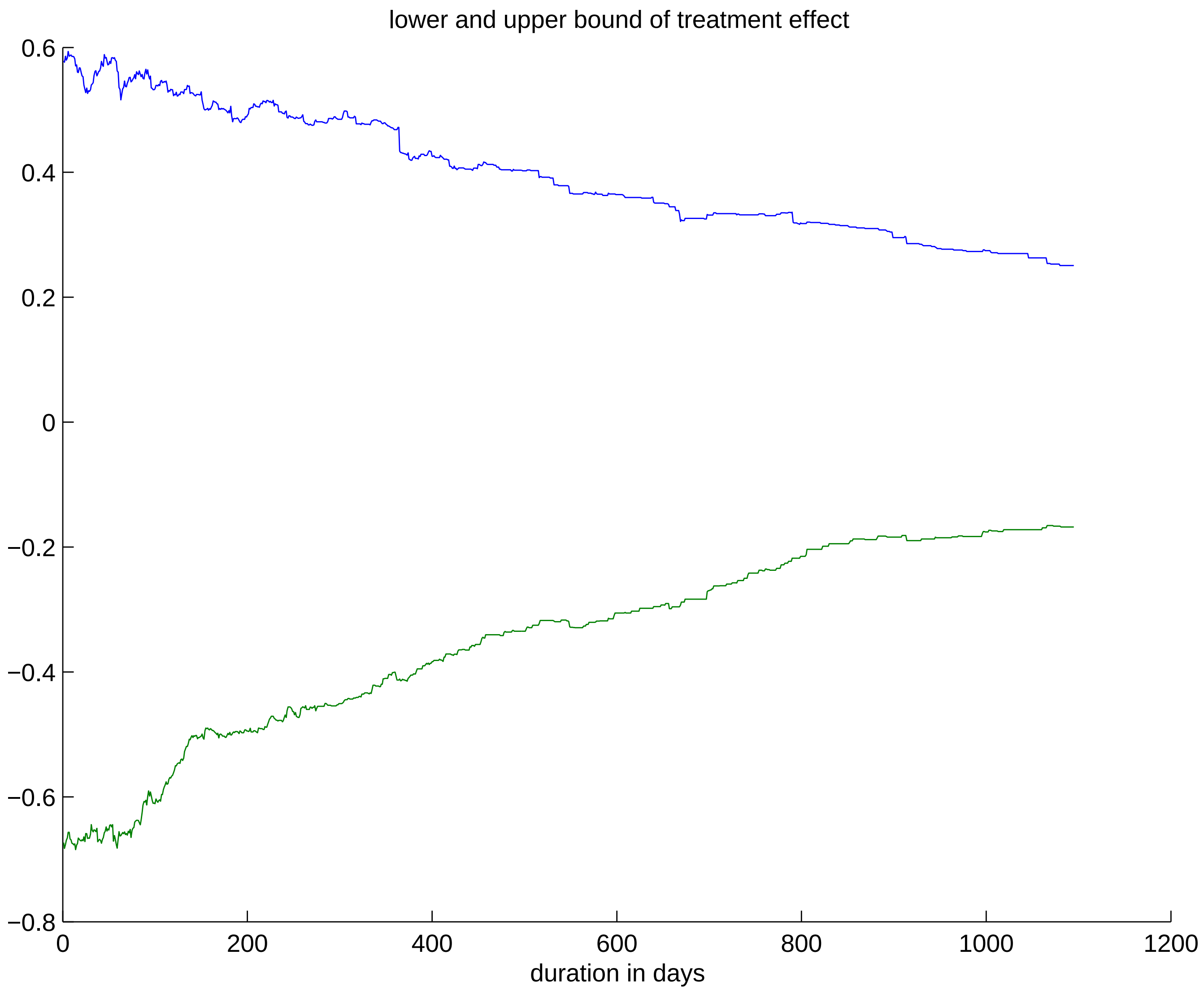

Figure 1. Without S1 and S2 (Male; Married)

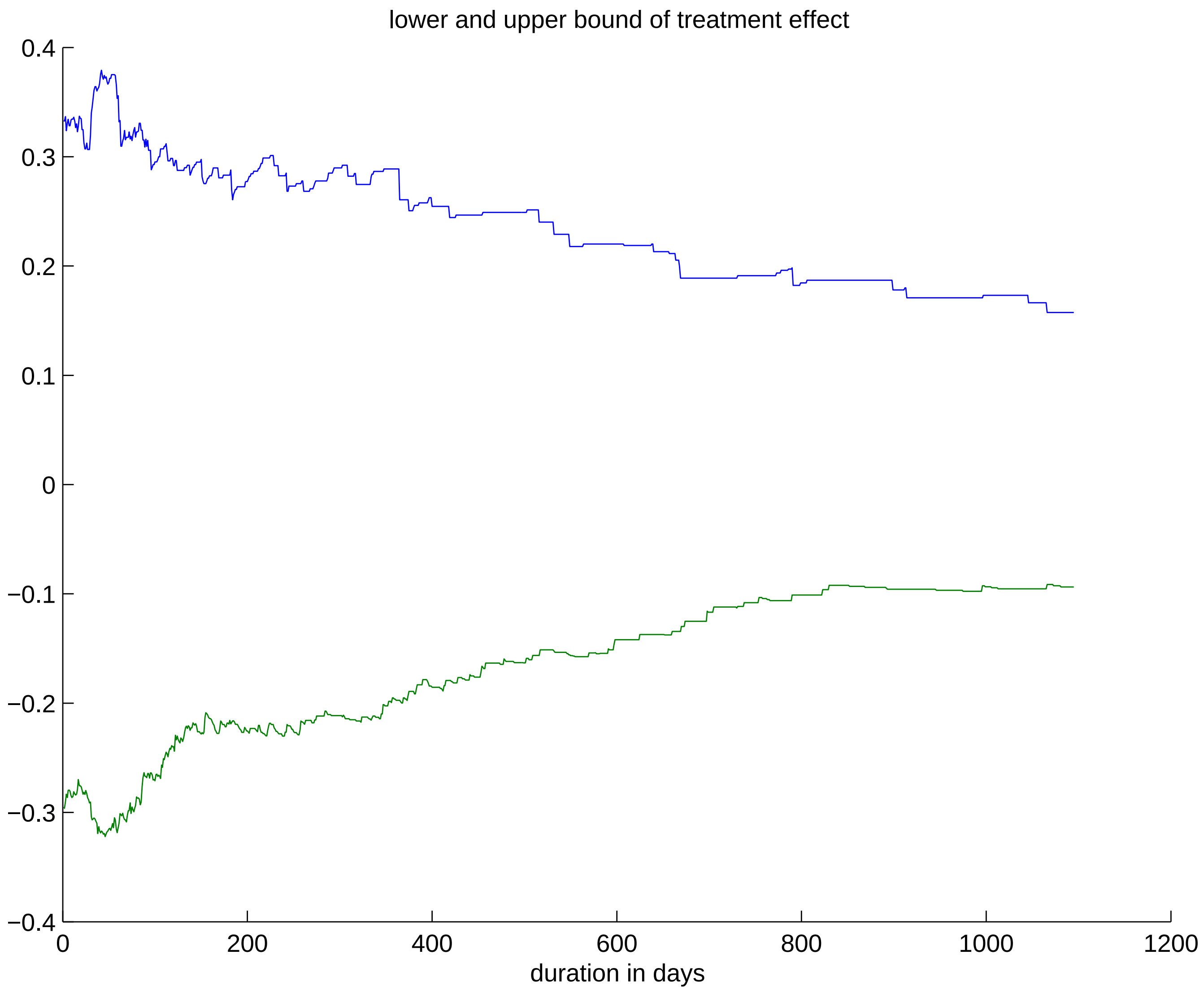

Figure 2. With only S1 (Male; Married)

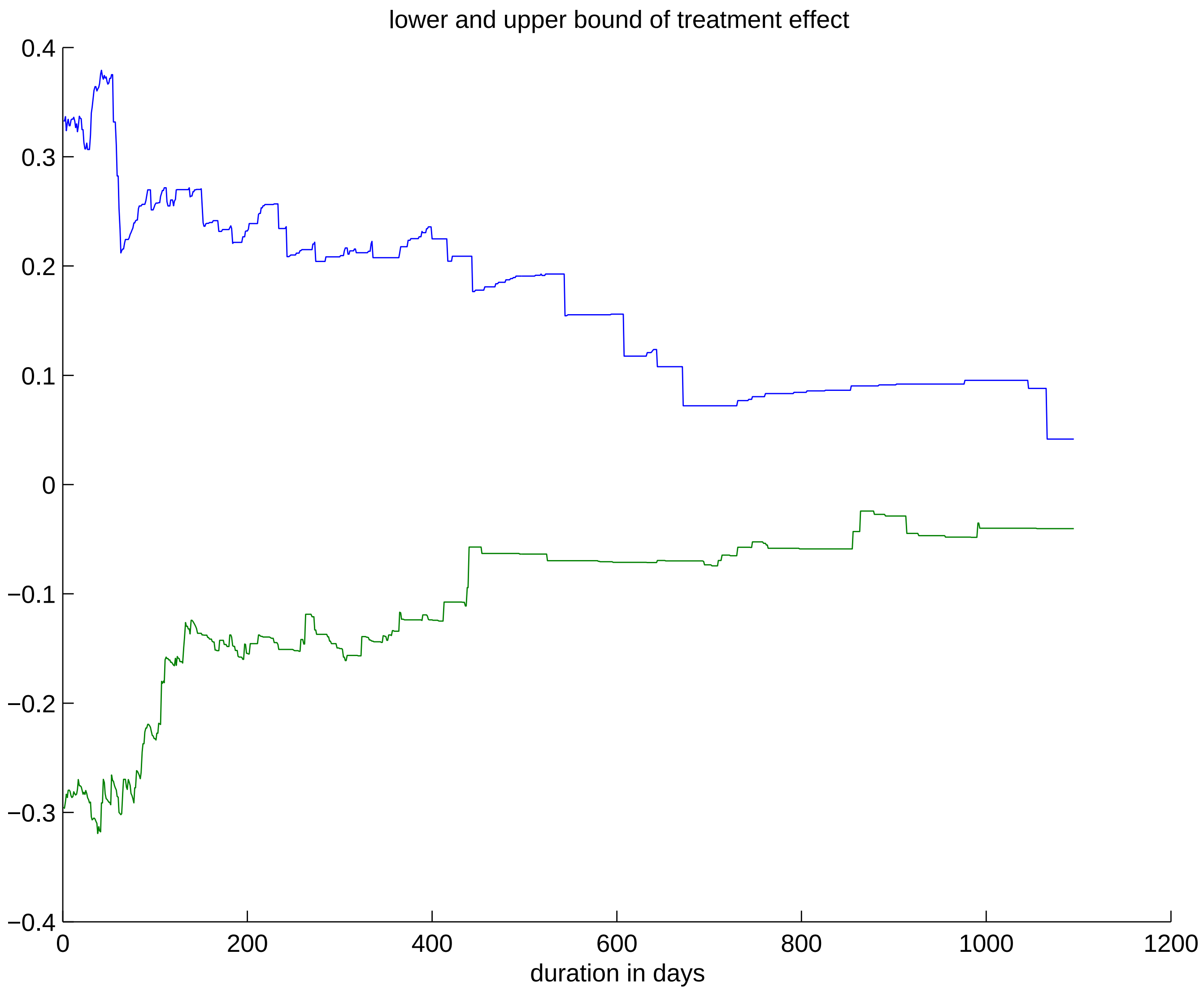

Figure 3. With S1 and S2 (Male; Married)

Bounding quantile treatment effects

• Notice (1) can be rewritten in terms of conditional quantile functions:

$$Q_1(\tau|x) \le Q(\tau|x) \le Q_2(\tau|x), \tag{5}$$

where $Q(\tau|x)$ is the τ-th quantile of Y conditional on $X = x$ and $Q_j(\tau|x)$ is the τ-th quantile of Y_j conditional on $X = x$ for $j = 1, 2$.

• Invoking difference-in-differences strategy to identify quantile treatment effects (see, for example, Athey and Imbens, 2002), we have

$$\begin{aligned}\Delta_Q(\tau|x) = {} & [Q(\tau|T=1, P=1, X=x) - Q(\tau|T=1, P=0, X=x)] \\ & - [Q(\tau|T=0, P=1, X=x) - Q(\tau|T=0, P=0, X=x)].\end{aligned}$$

Bounding quantile treatment effects (continued)

- Hence, using the empirical evidence alone, we have a bound for $\Delta_Q(\tau|x)$:

$$l_Q(\tau|x) = [Q_1(\tau|1,1,x) - Q_2(\tau|1,0,x)] - [Q_2(\tau|0,1,x) - Q_1(\tau|0,0,x)]$$

and

$$u_Q(\tau|x) = [Q_2(\tau|1,1,x) - Q_1(\tau|1,0,x)] - [Q_1(\tau|0,1,x) - Q_2(\tau|0,0,x)].$$

Bounding quantile treatment effects (continued)

- As before, we can make additional assumptions.
- **Assumption Q1:** $Q(\tau|0, p, x) \leq Q(\tau|1, p, x)$
- Note that Assumption Q1 is equivalent to Assumption S1 if Assumption Q1 holds for each τ.
- Under this, we have a tighter bound

$$\tilde{l}_Q(\tau|x) = \max[0, Q_1(\tau|1, 1, x) - Q_2(\tau|0, 1, x)] - [Q_2(\tau|1, 0, x) - Q_1(\tau|0, 0, x)]$$

and

$$\tilde{u}_Q(\tau|x) = [Q_2(\tau|1, 1, x) - Q_1(\tau|0, 1, x)] - \max[0, Q_1(\tau|1, 0, x) - Q_2(\tau|0, 0, x)].$$

Bounding quantile treatment effects (continued)

• **Assumption Q2:** $Q(\tau|t,p,x,v) \leq Q(\tau|t,p,x,v')$ **for** $v \leq v'$ **for any** τ**.**

• Equivalently, $F(y|t,p,x,v) \geq F(y|t,p,x,v')$ for $v \leq v'$, where F denotes a conditional CDF. Then

$$\max_{v' \geq v} F_2(y|t,p,x,v') \leq F(y|t,p,x,v) \leq \min_{v' \leq v} F_1(y|t,p,x,v') \tag{6}$$

for $t = 0, 1$ and $p = 0, 1$.

Bounding quantile treatment effects (continued)

- By the law of iterative expectations,

$$R_2(y|t,p,x) \leq F(y|t,p,x) \leq R_1(y|t,p,x), \tag{7}$$

where

$$R_2(y|t,p,x) = \sum_v \Pr(V = v|t,p,x)[\max_{v' \geq v} F_2(y|t,p,x,v')]$$

and

$$R_1(y|t,p,x) = \sum_v \Pr(V = v|t,p,x)[\min_{v' \leq v} F_1(y|t,p,x,v')].$$

Bounding quantile treatment effects (continued)

• Notice that for any two invertible functions $f(x)$ and $g(x)$, $f(x) \leq g(x)$ if and only if $f^{-1}(x) \geq g^{-1}(x)$. Using this,

$$R_1^{-1}(\tau|t,p,x) \leq Q(\tau|t,p,x) \leq R_2^{-1}(\tau|t,p,x),$$

where $R_j^{-1}(\cdot|t,p,x)$ is the inverse of $R_j(\cdot|t,p,x)$ given (t,p,x) for $j = 1, 2$.

Bounding quantile treatment effects (continued)

- Combining this with Assumption Q1 gives a tighter bound for $\Delta_Q(\tau|x)$:

$$\bar{l}_Q(\tau|x) = \max[\tilde{l}_Q(\tau|x),$$

$$\max[0, R_1^{-1}(\tau|1,1,x) - R_2^{-1}(\tau|0,1,x)] - [R_2^{-1}(\tau|1,0,x) - R_1^{-1}(\tau|0,0,x$$

and

$$\bar{u}_Q(\tau|x) = \min[\tilde{u}_Q(\tau|x),$$

$$[R_2^{-1}(\tau|1,1,x) - R_1^{-1}(\tau|0,1,x)] - \max[0, R_1^{-1}(\tau|1,0,x) - R_2^{-1}(\tau|0,0,x$$

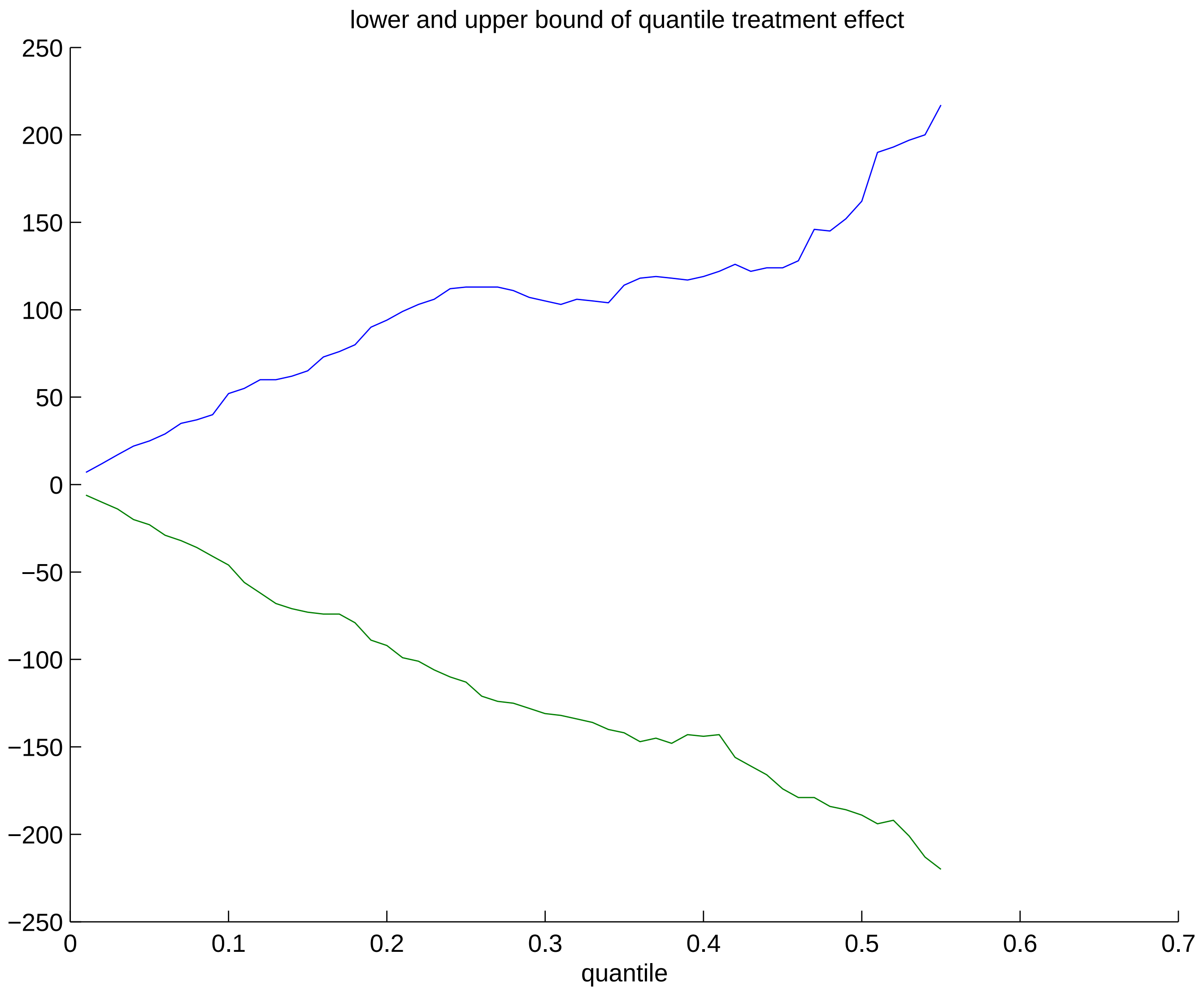

Figure 4. Without Q1 and Q2 (Male; Married)

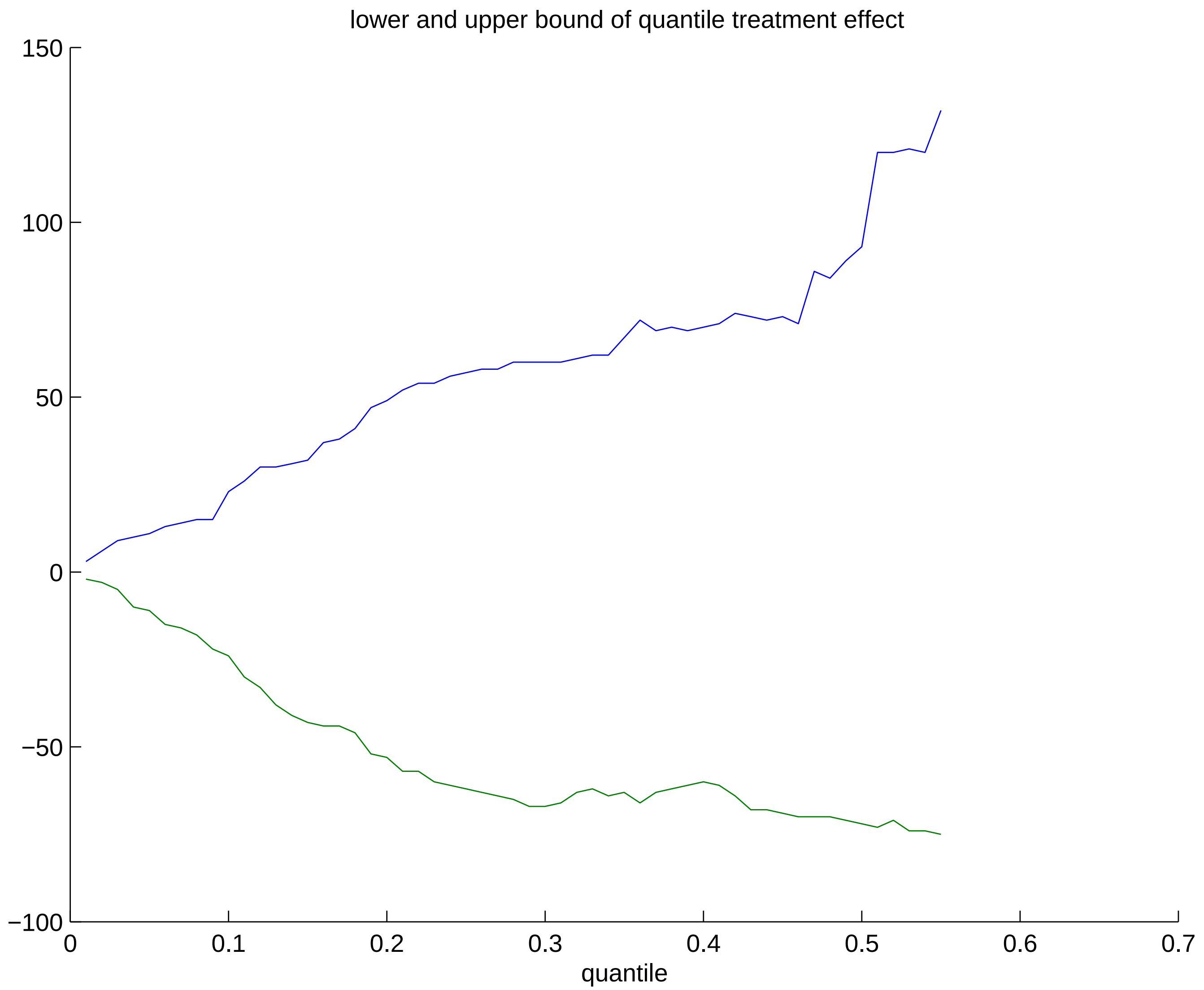

Figure 5. With only Q1 (Male; Married)

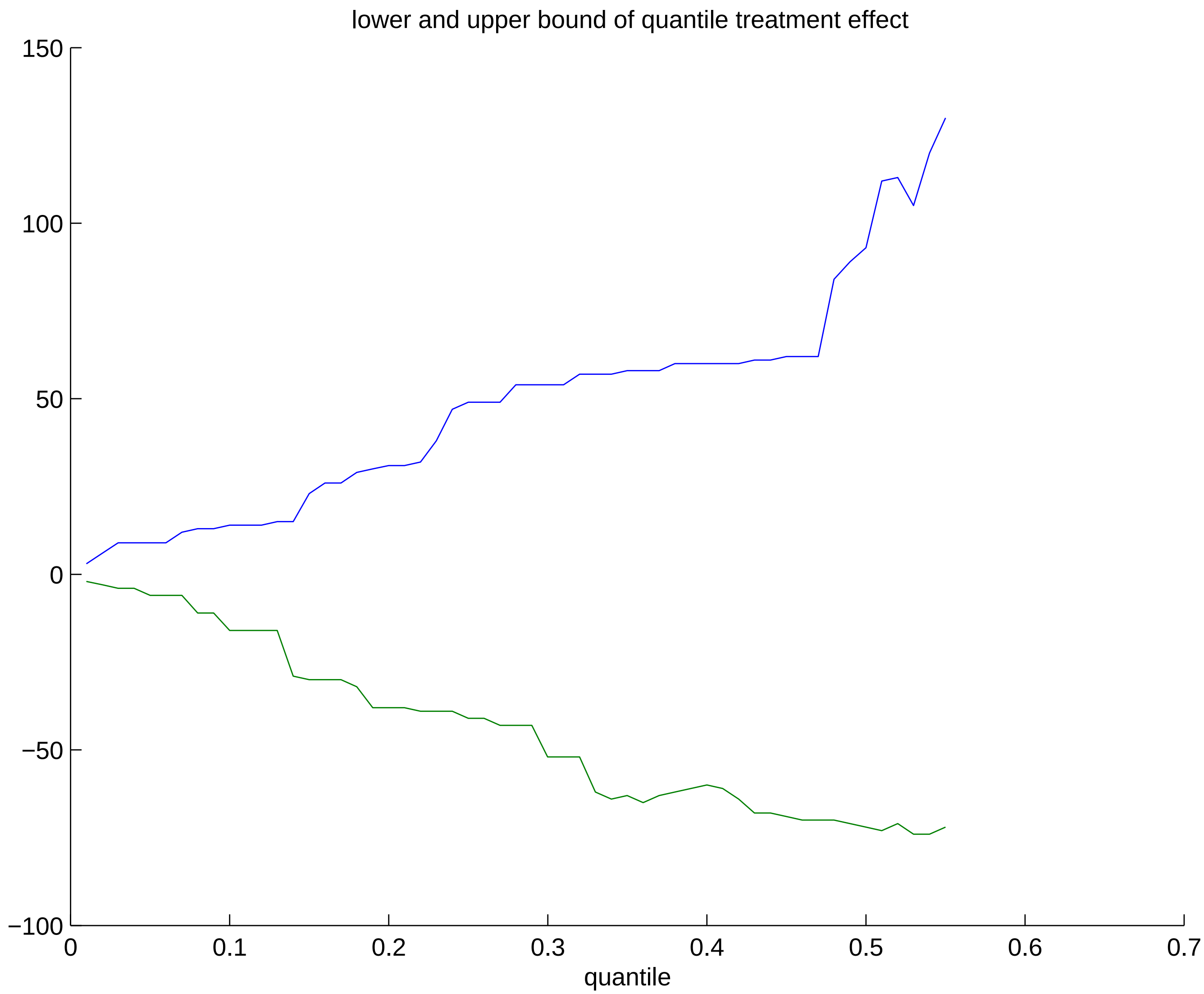

Figure 6. With Q1 and Q2 (Male; Married)

Conclusion and Future Work

- No evidence for supporting the claim that extensions of unemployment benefits increased unemployment durations.
- One may use a parametric duration model to further tighten the bounds.