Multiple Treatments

Oscar Mitnik
Juan Mejalenko

June 6, 2017

Office of Strategic Planning and Development Effectiveness

Introduction

- Typically we see cases where there is a treatment and a control group
- We usually compare treatment effects for both groups (ATE) or for those individuals who participated (ATT)
- If the treatment and control groups were chosen at random → we just compare the two groups (either by difference of means or with regressions)
- If assignment was not random → we use non-experimental methods (Diff-in-Diff, Matching) or quasi-experimental (IV, RD)

Continuous & multiple treatments

- However, many times there are programs where:
  - Treatment is **continuous**:
    - Treatment intensity: program lasts longer for some participants (entry/exit)
    - Participants receive different amounts of subsidies
  - There are **multiple treatments**:
    - Treatment 1: Information
    - Treatment 2: Subsidy
    - Treatment 3: Subsidy + Information
    - Control: No subsidy, no information
- When assignment to treatment is **random** then there is no problem → we just compare the results of each group
- But, if assignment to treatment is **not random** we have to deal with possible **selection bias** as in binary cases

Selection on observables

- When using regression to estimate treatment effects → we rely on "adjusting" regression with the observed variables
- Implicit assumption: **"selection on observables"**
  - All the variables that we include in the regression are enough to eliminate **selection bias**
- Another way of calling this assumption:
  - Unconfoundedness
  - Conditional Independence Assumption (CIA)
- Matching is an alternative (and attractive) method that relies on this assumption

Strategy usually used in binary case

- Imbens (2015) proposes following steps
- Stage I: Design
  - We do not use outcome (Y) data → focus only on (X,T) → typically (not always) used to estimate propensity score
  - Assess **overlap** → how similar are the distributions
  - If overlap is not good: we discard individuals that fail overlap → **Changes group for which we estimate ATT or ATE!**
- Stage II: Assess Unconfoundedness
  - Only use (X,T) data → but it is more useful when X includes lagged values of Y → "pseudo" outcomes X_p
- Stage III: Analysis
  - Estimate ATT or ATE
  - Only on individuals satisfying overlap, and assuming we believe unconfoundedness is a credible assumption

Non-experimental methods for continuous & multiple treatments I

- We are interested in the **average causal effect of some treatment on some outcome**
- **Treatment**: $T \in \mathcal{T}$ → for each treatment t there is a **potential outcome** $Y_i(t)$
- We are interested in $E\{Y_i(t)\}$, for all values of t
  - And in differences of the type $E\{Y_i(t) - Y_i(s)\}$
- We observe for each unit i
  - Random sample of size N
  - Treatment T_i
  - Outcome associated to treatment: $Y_i \equiv Y_i(T_i)$
  - Pre-treatment variables X_i

Non-experimental methods for continuous & multiple treatments II

- Key assumption: adjusting for pre-treatment differences in X_i solves problem of drawing causal inference
  → **unconfoundedness assumption**
- If $D_i(t) = 1$ is a dummy for $T_i = t$ (0 otherwise)
- **Weak unconfoundesdness**: $\boldsymbol{D(t) \perp Y(t)|X}$ for all $\mathbf{t} \in \mathcal{T}$
  → treatment independent of each potential outcome **|X**
- As binary case problem of **| X** →multi-dimensional
- Imbens (2000): **Generalized Propensity Score (GPS)**
  - Probability of being assigned to particular treatment, or to particular intensity of treatment conditional on covariates
  - **GPS**: $r(t,x) \equiv pr(T = t|X = x) = E\{D(t)|X = x\}$

Estimation

- Imbens (2000) proposes estimation in three steps (Hirano & Imbens, 2004 apply it to **continuous case**):
  1. Estimate GPS (i.e. multinomial logit)
  2. For all t ∈ $\mathcal{T}$ estimate:
     i. $\beta(t,r) \equiv E\{Y|T=t, r(\boldsymbol{T},X)=r\}$
     ii. $E\{Y(t)\} = E\{\beta(t, r(\boldsymbol{t},X))\}$

  } **Partial Mean Estimator**

- Alternatively, as in binary case a type of "matching" can be done by **inverse probability weighting (IPW)**

$$E\{Y(t)\} = \frac{YD(t)}{r(T,X)}$$

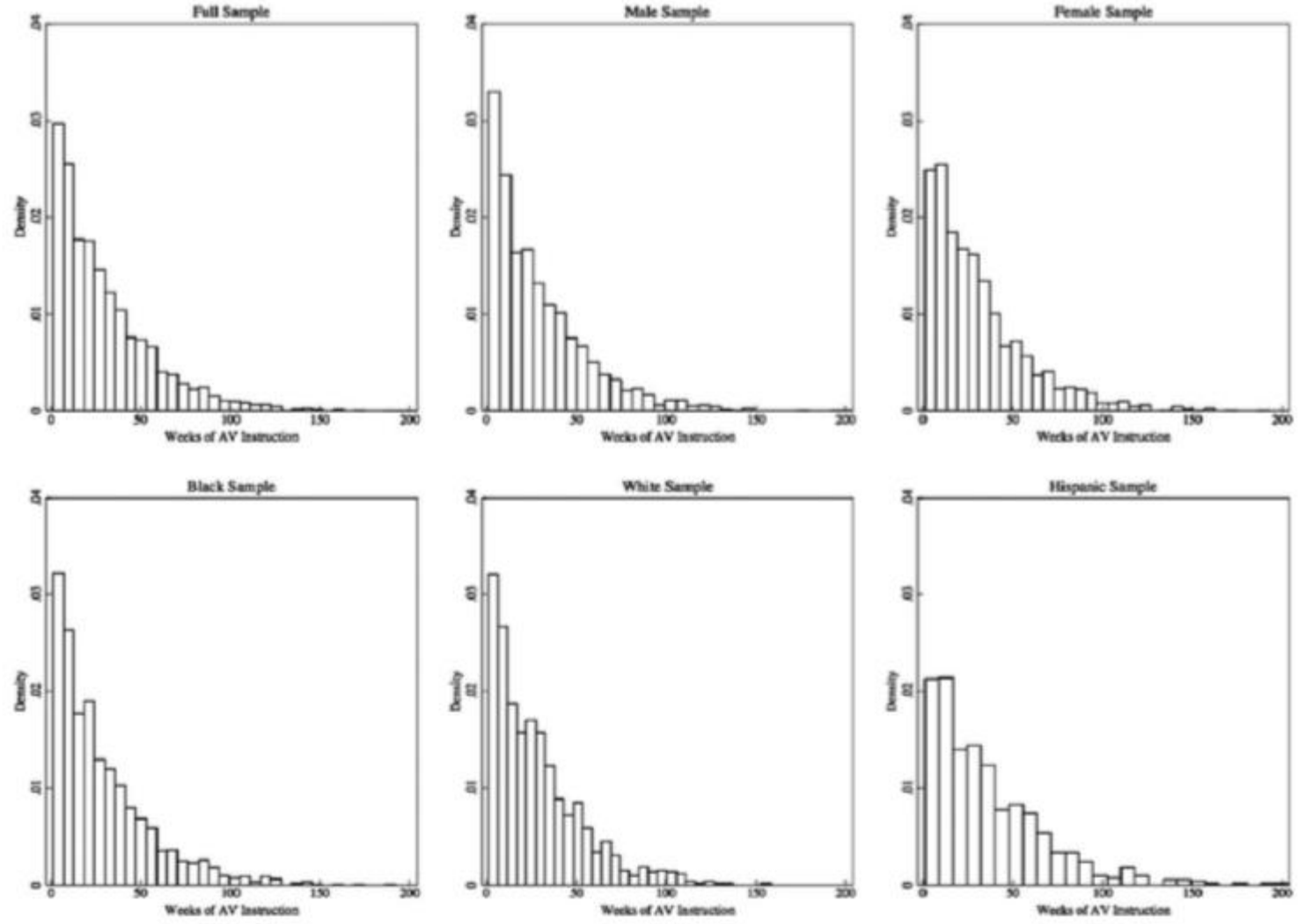
- Cattaneo (2010) shows how to do this **efficiently** with GMM
- IPW is implemented in **Stata**: `teffects aipw` command

Continuous treatment Example I

- Flores et al. (2011) analyze the effects of length of exposure to instruction in the Job Corps training program
- The JC program is a U.S. program oriented to young high school dropouts (ages 16 to 24)
- Approximately 60,000 participate per year at a cost of $1b
- Participants receive many services (including residing in the training centers, health services, and training)
- Duration of effective academic and vocational (AV) instruction varies by type of training, and by individual → different types of individuals end up receiving a different treatment length → selection bias?

Continuous treatment Example I

FIGURE 1.—HISTOGRAMS OF THE LENGTH OF EXPOSURE TO ACADEMIC AND VOCATIONAL INSTRUCTION IN JC FOR EACH DEMOGRAPHIC GROUP



Continuous treatment Example I

- Authors estimate the GPS for the continuous case
- As shown by Hirano & Imbens (2004) not enough to run an (inverse) weighted regression
- It is necessary to average this regression for different values of the GPS → Details in paper, beyond our focus
- The whole procedure can be done in Stata with the **`drf`** Stata command written by Bia, Flores-Lagunes, Flores & Mattei (2014) → estimates GPS, produces balancing tests, estimates "dose-response functions"

Continuous treatment Example I

- Balancing testes are more complicated
- They run a regression of T on GPS, GPS^2, GPS^3, Xs → they test that Xs have no explanatory power, after controlling for GPS

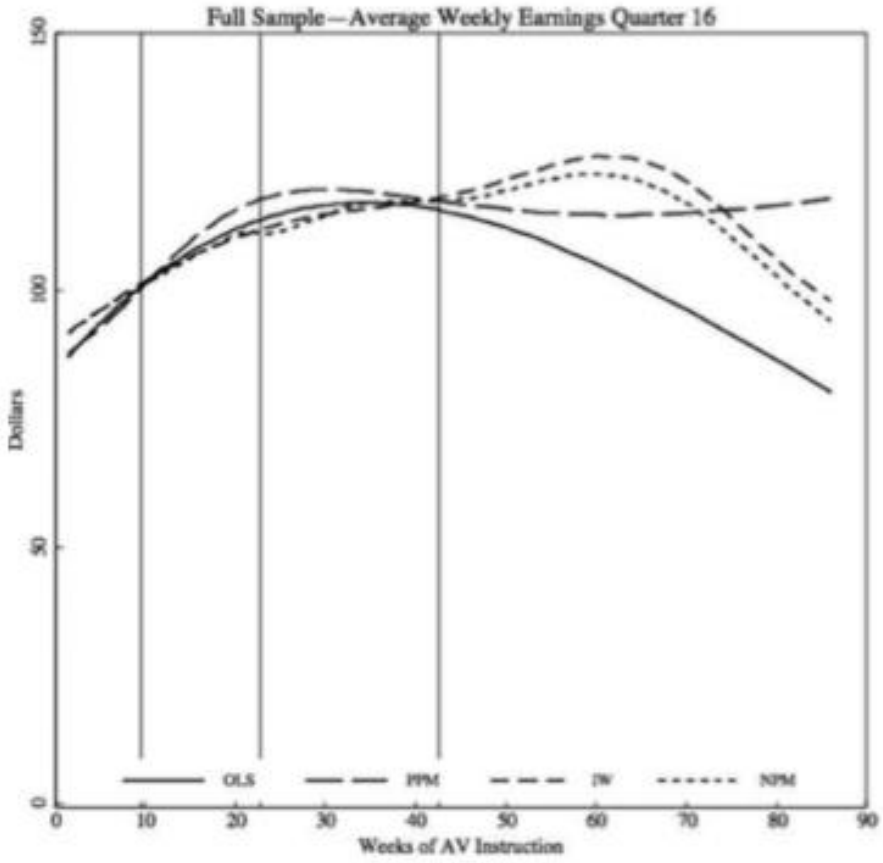
TABLE 2.—BALANCING TESTS

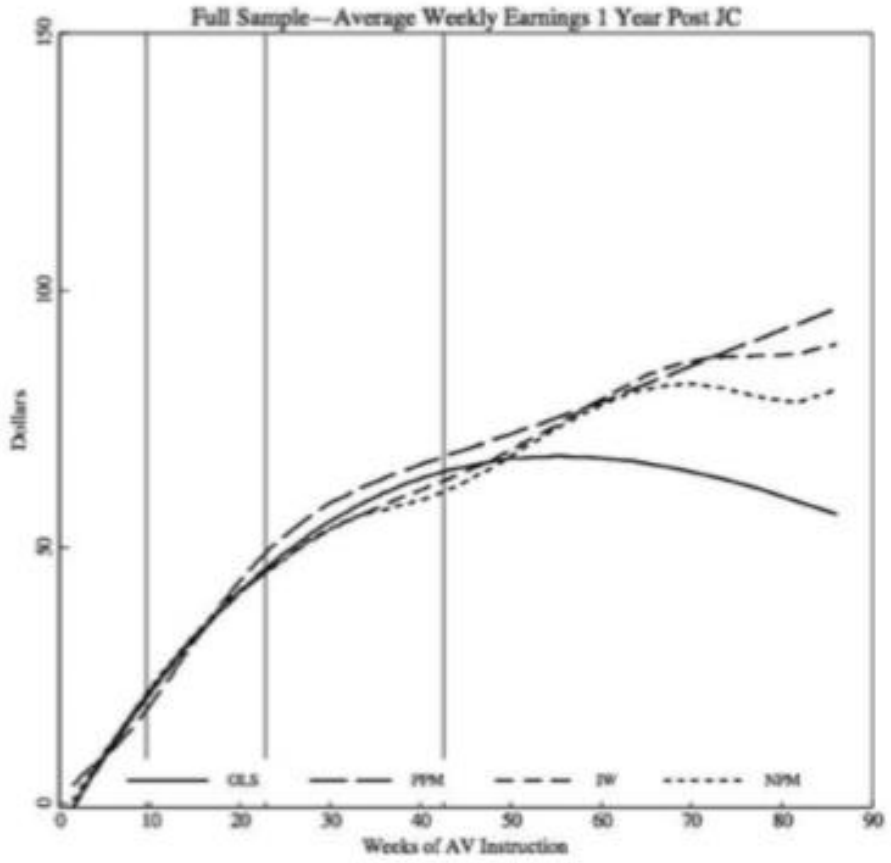
	Sample					
	Full	White	Black	Hispanic	Male	Female
Unrestricted model: T on GPS, GPS^2, GPS^3, and X's						
Test restriction that X's can be excluded from the unrestricted model						
Restricted log likelihood	−14,554	−3,088	−7,505	−1,846	−7,526	−5,946
Unrestricted log likelihood	−14,520	−3,069	−7,483	−1,828	−7,496	−5,924
Test statistic	69	38	44	35	61	43
p-value	1.00	1.00	1.00	1.00	1.00	1.00
Number of restrictions	405	253	243	209	333	322
Test restriction that GPS coefficients can be excluded from the unrestricted model						
Restricted log likelihood	−15,392	−3,186	−7,940	−1,871	−7,932	−6,229
Unrestricted log likelihood	−14,520	−3,069	−7,483	−1,828	−7,496	−5,924
Test statistic	1,745	235	915	87	873	610
p-value	0.00	0.00	0.00	0.00	0.00	0.00
Number of restrictions	3	3	3	3	3	3
Number of observations	3,524	726	1,830	404	1,825	1,407

Continuous treatment Example I

- They estimate “dose-response functions” that show the treatment effects of different weeks of AV instruction

FIGURE 2.—DOSE-RESPONSE FUNCTION ON TWO OUTCOMES (IN DIFFERENCES) USING ALTERNATIVE ESTIMATORS IN THE FULL SAMPLE

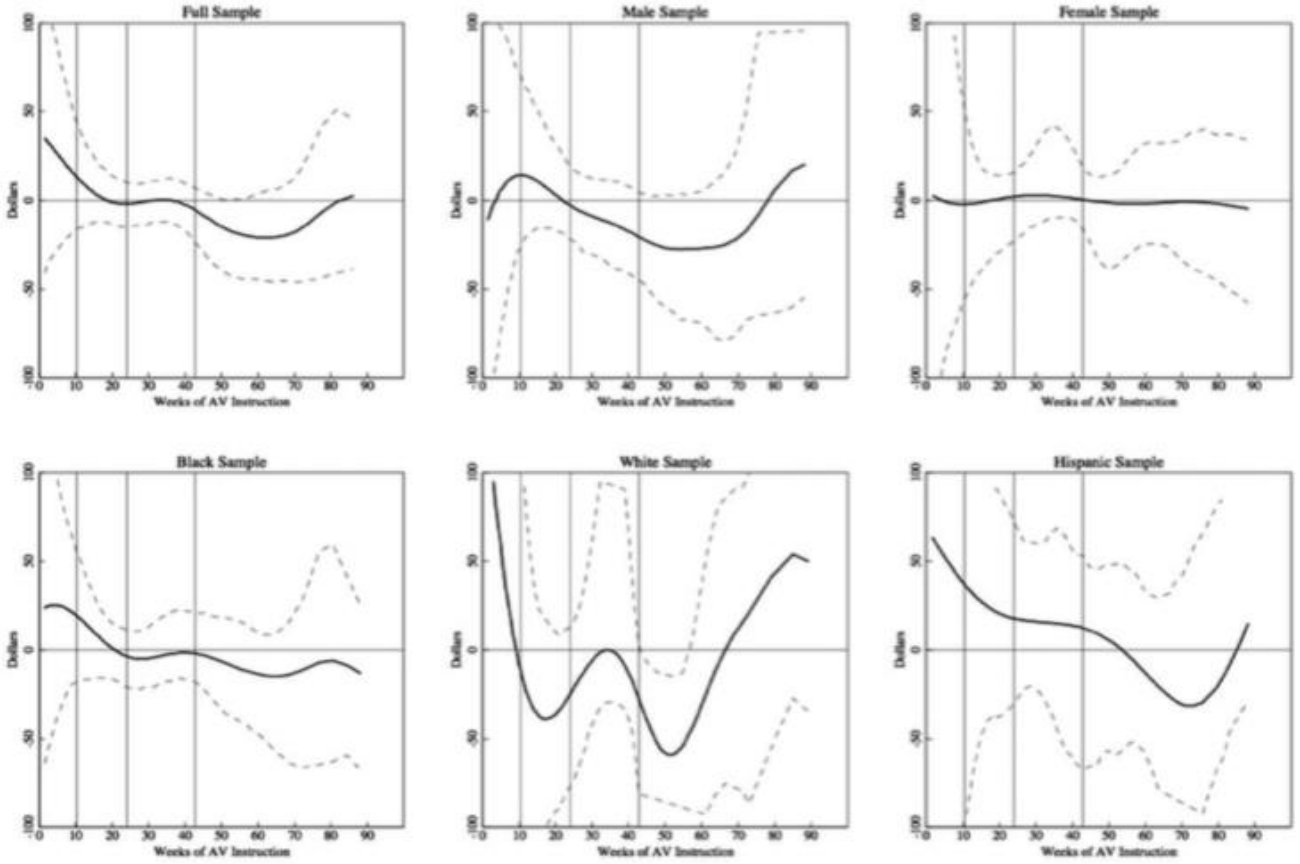




Continuous treatment Example I

- They validate results by looking at “placebo outcomes” (here in “derivative form”, i.e. changes)

FIGURE 5.—ESTIMATED DERIVATIVE OF THE DOSE-RESPONSE FUNCTION ON A PLACEBO OUTCOME (TOTAL EARNINGS IN THE FIRST THREE QUARTERS OF THE YEAR PRIOR TO RANDOMIZATION) USING THE IW ESTIMATOR



Intensity of treatment for participants

- A situation that is not uncommon is to have data only on program participants, but with different lengths of exposure
- **In this case we cannot estimate a treatment effect where the counterfactual is not participating in the program**
- But, it can still be interesting to understand what are the effects of length of exposure for participants
- **Participants are most likely very different from non-participants, so we need to be careful in interpreting these analyses!!!**

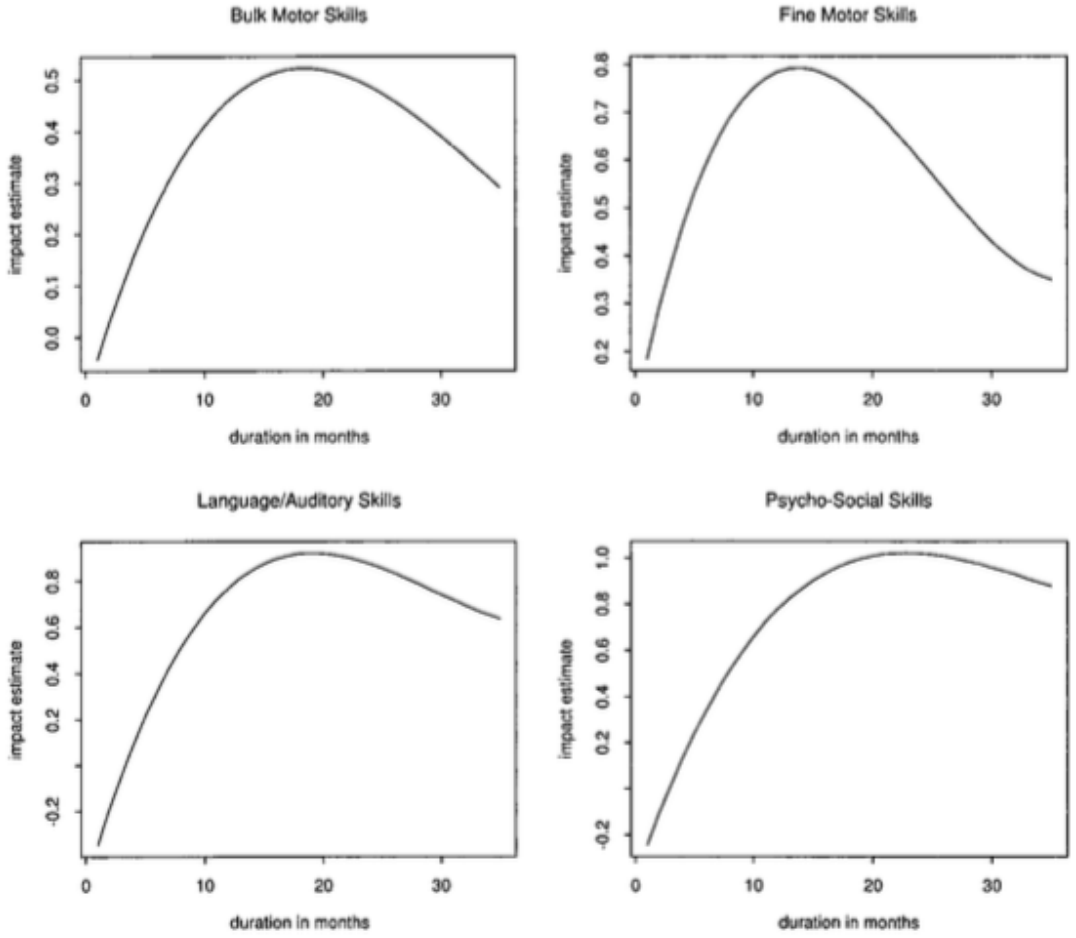
Continuous treatment Example II

- Behrman, Cheng & Todd (2004) evaluate the “Proyecto Integral de Desarrollo Infantil (PIDI)” in Bolivia
- The program provides daycare, nutritional, and educational services to children 6 months to 6 years who live in poor, predominantly urban areas
- The authors do not rely on the GPS but do use a version of matching, and rely on a **marginal impact estimator**
- They also check balancing of covariates and impose a common support condition

Continuous treatment Example II

- Note how all curves start at values >0!!!

FIGURE 4.—ESTIMATED PROGRAM IMPACTS FROM CROSS-SECTIONAL, CUBIC-IN-DURATION MODEL, SAMPLES *P* AND *B*



Multiple (multi-valued) treatments Example I

- Cattaneo (2010) analyzes the effect of maternal smoking intensity during pregnancy on birth weight
- He estimates marginal **mean** and **quantile** treatment effects → i.e. not only the average effect, but also showing (for example) the median effect (or any other quantile)
- He uses a version of inverse probability weighting, and also proposes the Efficient Influence Function Estimator EIFE) that is **efficient** (minimizes variance)
- These estimators (and others) are coded in Stata 13 & 14 → see **`teffects multivalued`** in Stata help

Multi-valued treatments Example I

- Cattaneo (2010) findings
- Intensity of treatment matters
- Not only at the mean
- Also at different quantiles

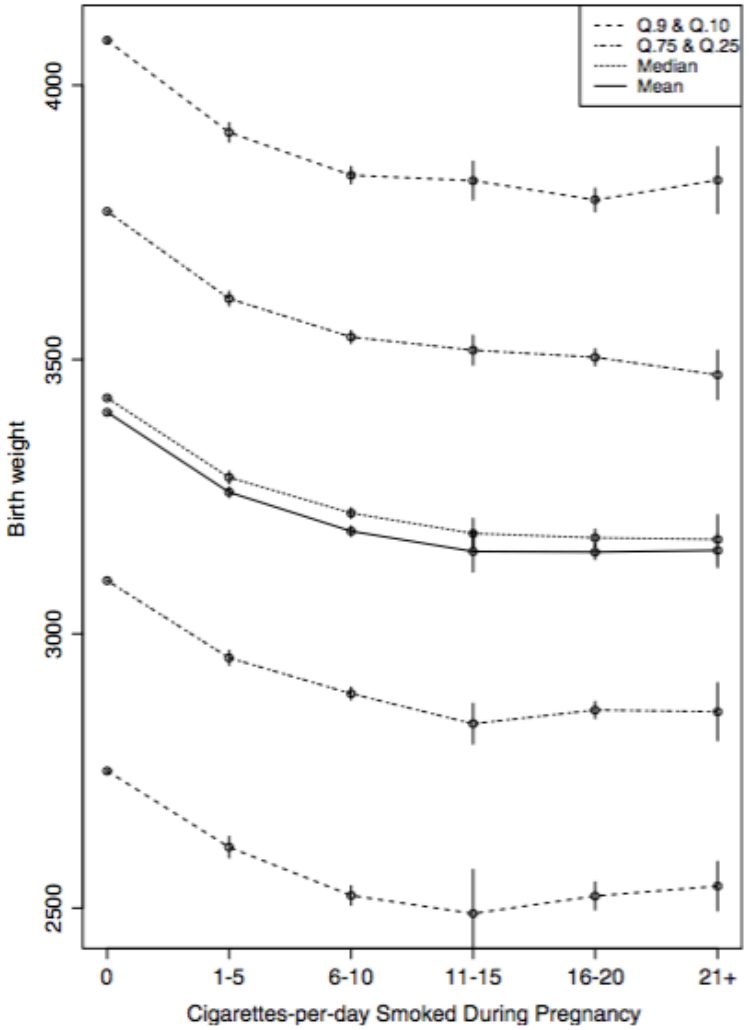


Fig. 1. Effect of maternal smoking intensity on birth weight (5-cigarette bins).

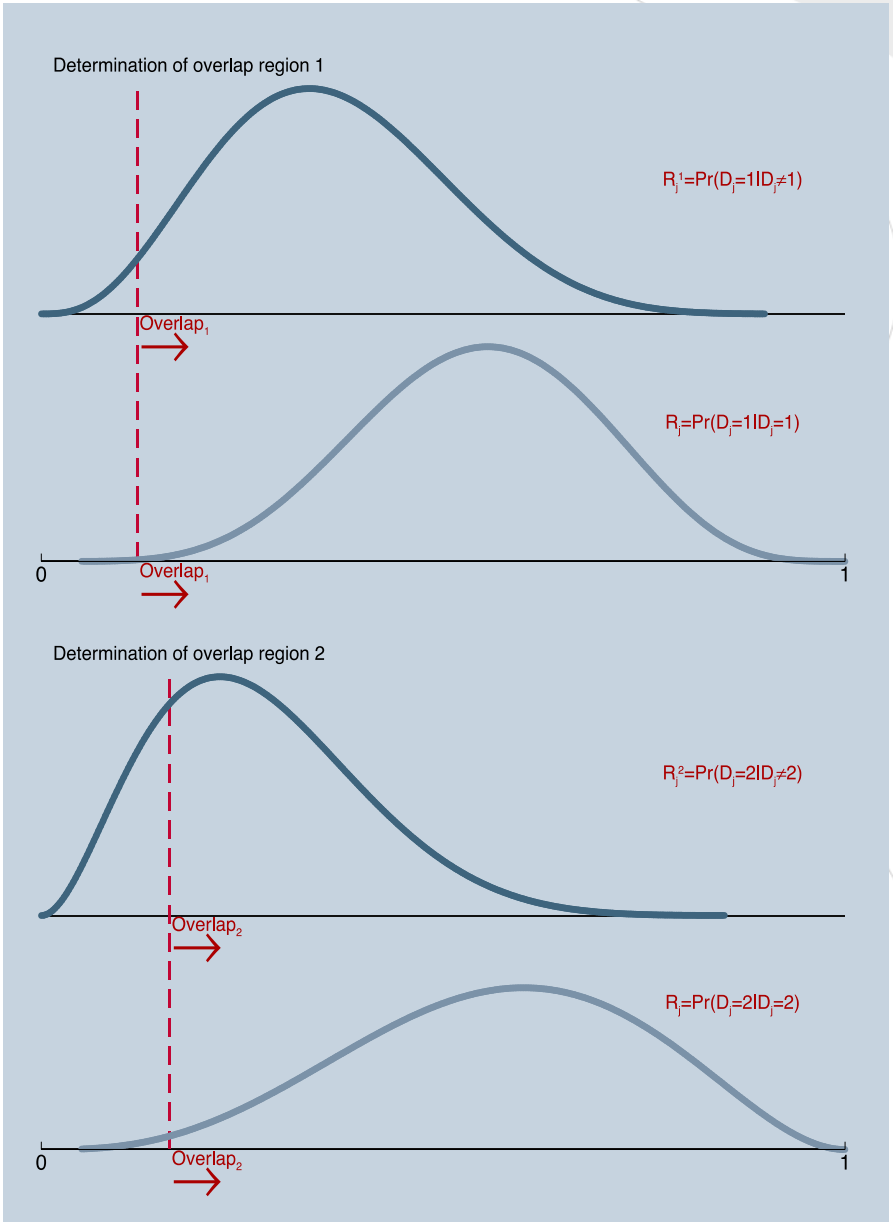
Multi-valued treatments Example II

- Flores & Mitnik (2013) rely on multi-valued treatment estimators to show how they can be used to compare treatments across different labor markets
- They analyze a training program in the U.S. in five different locations
- But, only use **control** individuals to make the point that selection bias can be eliminated through the use of these estimators
- Also, they show the importance of imposing **a common support condition** that makes individuals comparable across all five locations

Multi-valued treatments Example II

- How do you impose overlap?
- Let $r^{d}_{q,\{j\in A\}}$ be the q-th quantile of GPS distribution of sub-sample A
- Overlap_d = Overlap region **in site d** is given by sub-sample of individuals with GPS > $\max(r^{d}_{q,\{j:\, D_j=d\}},\, r^{d}_{q,\{j:\, D_j\neq d\}})$
- Overlap or common support region:

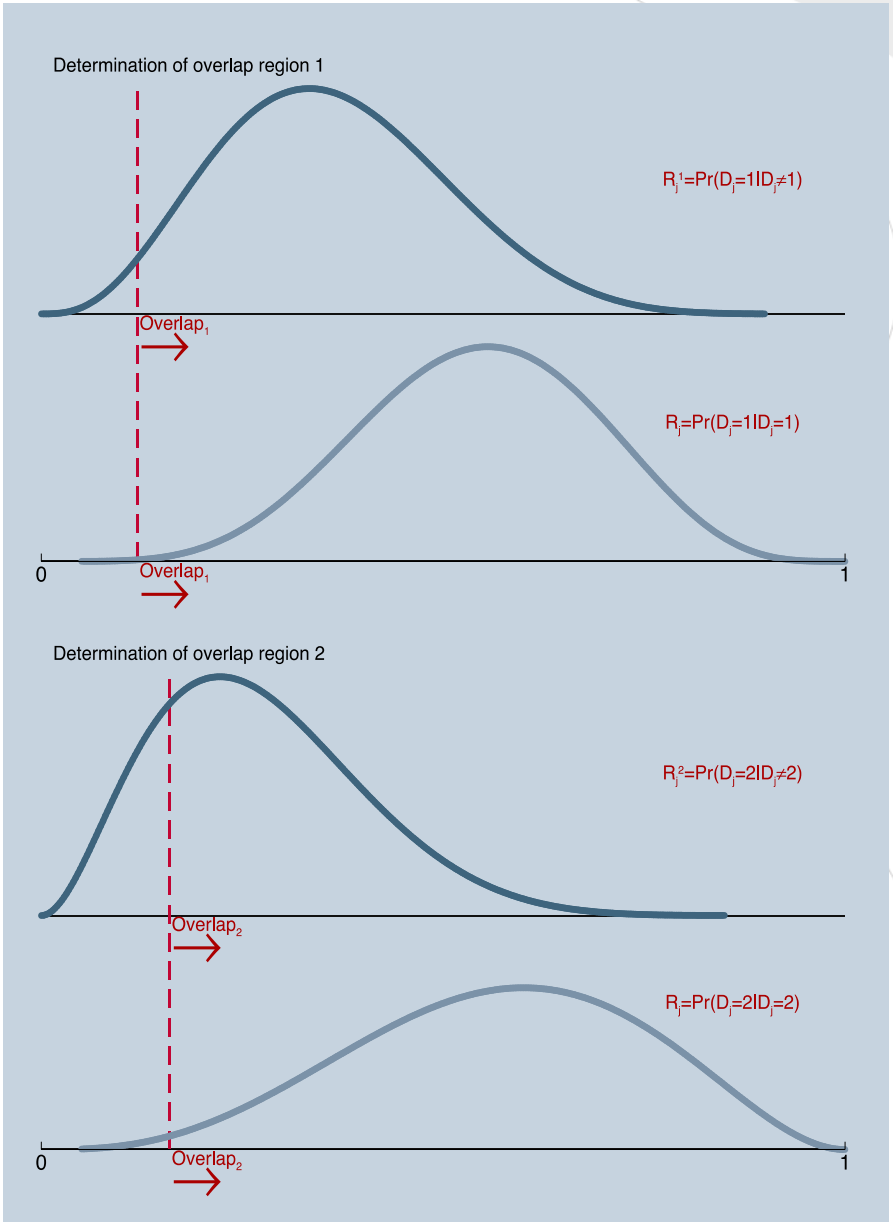
$$Overlap = \bigcap_{d=1}^{k} Overlap_d$$



Multi-valued treatments Example II

Note that:

- GPS of one treatment vs. all other treatments together are compared
- Only lower tails of the distributions are compared
  - For each individual in a particular treatment, looking for comparable individuals
- No set value for q
  - In paper q=0.0025 is used
  - Other values of q tried



Multi-valued treatments Example II

- Overlap (common support) matters a lot in this case

C. Overlap

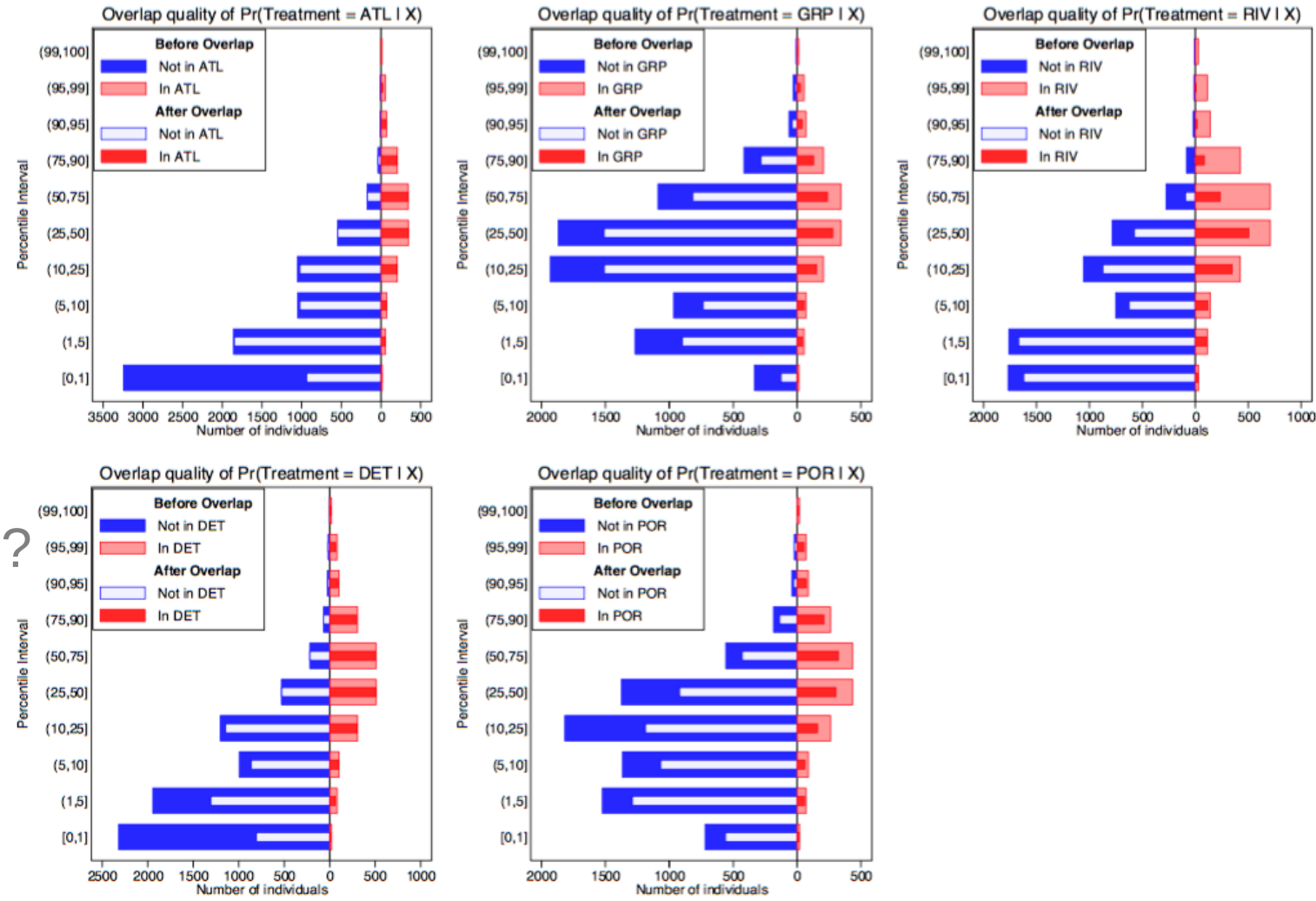
	ATL	DET	GRP	POR	RIV	Total
Number of observations	1,372	2,037	1,374	1,740	2,828	9,351
Number of observations after imposing overlap	1,299	1,933	969	1,263	1,437	6,901
Percentage observations dropped due to overlap	5.3%	5.1%	29.5%	27.4%	49.2%	26.2%

- Overlap has to be satisfied **simultaneously** in all 5 locations
- Based on GPS → Pr(belonging to a particular location)
- So in this case there are 5 different probabilities!
- This guarantees individuals are comparable in all locations (treatments)

Multi-valued treatments Example II

- Also important to check overlap quality
- How many units are comparable?

Figure 1. Overlap quality based on GPS distribution



Notes:
Outside wide bars and inside thin bars display the number of individuals whose GPS fall in a percentile interval, before and after imposing overlap, respectively.
Percentile intervals in each site are constructed using the GPS distribution, before imposing overlap, for individuals in that site only.

Aplicación práctica

mteffects (beta): programa para la aplicación práctica de múltiples tratamientos

- Selección de individuos comparables entre múltiples tratamientos.
- Análisis de los individuos seleccionados. Balanceo, Comparativa de las distribuciones (histogramas, densidades de Kernel).
- Diversidad de estimadores para las medias:
  - ✓ Sin ajustes (antes y después de overlap): RAW_NOV, RAW_OV
  - ✓ Media parcial basada en regresión lineal: PM_X_NOV, PM_X_OV
  - ✓ Media parcial basada en GPS (paramétrica y no paramétrica): PM_GPS_PAR, PM_GPS_NPR
  - ✓ Ponderación por inversa de probabilidad: IPW_NOX, IPW_X
- Simulaciones para establecer una base de comparación entre los tratamientos

Aplicación práctica

Caso 1: Tratados vs Controles

Caso 2: Múltiples ciudades. Buscar individuos comparables **entre ciudades**

Caso 3: Efectos del tratamiento **dentro** de las ciudades

Caso 4: Efectos del tratamiento **dentro** de las ciudades
→ con especificaciones diferentes para cada ciudad

Caso 5 (si hay tiempo): Comparando controles entre las ciudades**.**
Balanceo. Distribuciones de Kernel. Regiones de (no)overlap

Ejemplo usando:

- Attanasio, O., Kugler, A., & Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, *3*(3), 188-220.

Datos: https://www.aeaweb.org/aej/app/data/2009-0168_data.zip

Caso 1: Tratados vs Controles

Analisis sencillo.

Medias para cada grupo.

Diferencia de las medias**.**

Estimadores a utilizar

- Media sin ajustar (después de overlap): RAW_OV
- Media parcial basada en regresión lineal: PM_X_NOV
- Ponderación por inversa de la probabilidad: IPW_X

Caso 1: Tratados vs Controles

Preparamos la base de datos

```
*** Cargamos la base de datos
use "C:\Users\jmejalenko\Desktop\AEJApp-20090168_data.dta" , clear
keep if dcontinue==1        // Nos quedamos unicamente con los individuos con
                            // informacion en el Seguimiento
```

Generamos un identificador por curso (necesario para la estimación)

```
egen course = group(codigo_curs codigo_ecap city) // Generamos un identificador
                // por curso, al nivel que se hizo la randomizacion
```

Definimos una macro local que contiene las variables de la línea de base que vamos a utilizar

```
*Definimos una macro local que contenga las variables en la linea de base
local bl_cov "salary_04 profit_04 tenure_04 days_04"
local bl_cov "`bl_cov' hours_04 contract_04 dformal_04 ldays_04"
local bl_cov "`bl_cov' lhours_04 pempl_04 age_lb dmarried_lb educ_lb"
```

Caso 1: Tratados vs Controles

Sintaxis base:

mteffects yvar , tvar( ) xvar( ) xgps( )

estimators( )

RAW_NOV, RAW_OV, PM_X_NOV, PM_X_OV,
PM_GPS_PAR, PM_GPS_NPR, IPW_NOX, IPW_X

treatment_lbls( )

Por default el programa corre un bootstrap de 200 replicaciones y los intervalos de confianza reportados al 95%. “rep()” y “level()”

Caso 1: Tratados vs Controles

Definimos variables macro para facilitar la visualización de la programación

```
*Macro con nuestra(s) variable dependiente(s)
local yvar "dformal_06"
*Macro con nuestra variable de tratamiento
local tvar "select"
*Macro con nuestras co-variables
local xvar "`bl_cov' i.course"
*Macro con las variables a utilizar en la estimacion del GPS
local xgps "`bl_cov'"
*Macro con los estimadores a utilizar
local estimators "RAW_OV PM_X_NOV IPW_X"
*Macro con los nombres de los tratamientos
local treatment_lbls "CONTROLS TREATED"
```

```
set seed 123456
mteffects `yvar' , tvar(`tvar')  xvar(`xvar') xgps(`xgps')          ///
            estimators(`estimators')                                ///
            treatment_lbls(`treatment_lbls')
```

Caso 1: Tratados vs Controles

Bootstrap results — Number of obs = 1,434; Replications = 200

Outcome variable: **dformal_06**

	Observed Coef.	Bias	Bootstrap Std. Err.	[95% Conf.	Interval]	
RAW_OV						
CONTROLS	.31653544	.001103	.01745186	.2769516	.349182	(P)
TREATED	.3709273	-.000216	.01706071	.3310124	.4078627	(P)
PM_X_NOV						
CONTROLS	.31909395	.0032512	.02276219	.2739274	.3673493	(P)
TREATED	.36849154	-.0015089	.01899902	.3273231	.4012094	(P)
IPW_X						
CONTROLS	.31821335	.0025195	.02166966	.2740694	.3612262	(P)
TREATED	.36776603	-.0017502	.01982395	.325104	.4012494	(P)

(P) percentile confidence interval

Caso 2: Múltiples ciudades

Queremos observar los efectos en individuos que sean comparables entre ciudades. Generamos el identificador de

```
/*
Etiquetas de las ciudades segun el valor que toman en la variable original
BARRANQUILLA = 1    BOGOTA = 2      BUCARAMANGA = 3     CALI    = 4
CARTAGENA = 5       MANIZALES = 6   MEDELLIN = 7                       */

gen     ciudad = 1 if city==2
replace ciudad = 2 if city==4
replace ciudad = 3 if city==7
replace ciudad = 4 if inlist(city, 1, 3, 5, 6)
```

Generamos una variable de tratamiento que contenga un valor para cada grupo que nos interese analizar

```
*Generamos la variable de tratamiento
egen tvar = group(select ciudad)
tab  tvar select
```

Caso 2: Múltiples ciudades

Volvemos a definir las macros necesarias

```
*Definimos las macros como anteriormente
local bl_cov "salary_04 profit_04 tenure_04 days_04"
local bl_cov "`bl_cov' hours_04 contract_04 dformal_04 ldays_04"
local bl_cov "`bl_cov' lhours_04 pempl_04 age_lb dmarried_lb educ_lb"

*Definimos macros con nuestra(s) variable(s) dependiente(s), covariables para
* la estimacion del gps y las regresiones
local yvar "dformal_06"
local tvar "tvar"
local xvar "`bl_cov'"
local xgps "`bl_cov'"

*Definimos los estimadores a utilizar y los nombres de los tratamientos
local estimators     "PM_X_NOV IPW_X"
local treatment_lbls "Bogota_C Cali_C Medellin_C Other_C"
local treatment_lbls "`treatment_lbls' Bogota_T Cali_T Medellin_T Other_T"
```

Caso 2: Múltiples ciudades

Llamamos al comando

```
set seed 123456
mteffects `yvar' , tvar(`tvar') xvar(`xvar') xgps(`xgps')                ///
          estimators(`estimators')                                       ///
          treatment_lbls("`treatment_lbls'")                             ///
          equation(#1%-#2%; #5%-#1%)
```

Nos permite realizar cómputos especiales reportados por el programa

Por ejemplo: diferencia entre los controles de Bogotá y Cali, o bien el efecto del tratamiento en Bogotá (Tratados - Controles)

Caso 2: Múltiples ciudades

	Observed Coef.	Bias	Bootstrap Std. Err.	[95% Conf.	Interval]	
PM_X_NOV						
Bogota_C	.3553111	.0018618	.03493103	.2923664	.4286339	(P)
Cali_C	.30596889	-.0036778	.04635415	.2076135	.4099092	(P)
Medellin_C	.31867427	.0044569	.03749606	.2502027	.3950747	(P)
Other_C	.27478694	.0016558	.03187954	.2066007	.3320753	(P)
Bogota_T	.42401556	.00195	.02922266	.3760134	.4856592	(P)
Cali_T	.36780572	-.0053266	.04364333	.277103	.4511535	(P)
Medellin_T	.36164645	.0009002	.04166647	.2864189	.4408595	(P)
Other_T	.31850559	-.0008908	.02859392	.2587556	.3692152	(P)
IPW_X						
Bogota_C	.36029295	.0061252	.04268033	.2929954	.4528752	(P)
Cali_C	.31438372	.0106535	.05594983	.2174356	.4402636	(P)
Medellin_C	.36014252	.0032864	.05321061	.2675115	.4612429	(P)
Other_C	.31669258	-.0103686	.0472014	.2011141	.3929699	(P)
Bogota_T	.44388858	.0021447	.04023667	.368533	.5251092	(P)
Cali_T	.354313	.0106046	.05119814	.2699943	.4651674	(P)
Medellin_T	.36845049	-.0111758	.05701742	.2399254	.4689399	(P)
Other_T	.34870227	-.000718	.04329046	.2609708	.4278342	(P)

Caso 2: Múltiples ciudades

Bootstrap results — Number of obs = **1,434**; Replications = **200**

Outcome variable: **dformal_06**

Special equations
Eqn1 : **#1 - #2**
Eqn2 : **#5 - #1**

	Observed Coef.	Bias	Bootstrap Std. Err.	[95% Conf.	Interval]	
(...)			(...)			
Eqn1						
PM_X_NOV	**.04934221**	**.0055396**	**.0632396**	**-.0618411**	**.1852592**	(P)
IPW_X	**.06870445**	**.0000883**	**.0429792**	**-.0110338**	**.1581266**	(P)
Eqn2						
PM_X_NOV	**.04590923**	**-.0045284**	**.0702101**	**-.0894321**	**.196269**	(P)
IPW_X	**.08359563**	**-.0039804**	**.05191273**	**-.025494**	**.1794194**	(P)

(P) percentile confidence interval

Caso 3: Efectos del tratamiento dentro de las ciudades

Vamos a calcular los efectos del tratamiento **dentro de cada ciudad**.

Incorporamos la opción “tvardif()”.

- Nos permite otra variable de tratamiento para calcular diferencias dentro de cada nivel de nuestra variable dependiente.
- Facilita el reporte y da flexibilidad
- (lo veremos en el “Caso 4”)

Mismas macros de yvar, gps, xvar

```
*Local macro w/outcome(s), macro w/treatment variable, & for the gps estimation
local yvar "dformal_06"
local xgps "`bl_cov'"
local xvar "`bl_cov"
```

Caso 3: Efectos del tratamiento dentro de las ciudades

En lugar de generar un valor para Tratados y otro para Controles en cada ciudad, solo identificamos las ciudades de nuestro interés.

```
gen     tvar = 1 if city==2
replace tvar = 2 if city==4
replace tvar = 3 if city==7
replace tvar = 4 if inlist(city, 1, 3, 5, 6)
local   tvar "tvar"
```

Indicamos al programa cual es la variable que identifica los tratados y controles

```
*Macro para indicar cual es la variable que marca la distincion entre T y C
local tvardif "select"
```

Solamente le daremos nombre a las ciudades

```
*Macro que indica los nombres de las ciudades
local treatment_lbls "Bogota Cali Medellin Other"
```

Caso 3: Efectos del tratamiento dentro de las ciudades

Llamamos al comando

```
set seed 123456
mteffects `yvar' , tvar(`tvar') xvar(`xvar') xgps(`xgps')                /
            estimators(`estimators')                                     /
            treatment_lbls(`treatment_lbls')                             /
            equation(#2%-#4%)   tvardif(`tvardif')
```

Opción agregada identifica quién es tratado y control

Caso 3: Efectos del tratamiento dentro de las ciudades

Special equations
Eqn1 : **#2 - #4**

	Observed Coef.	Bias	Bootstrap Std. Err.	[95% Conf.	Interval]	
PM_X_NOV						
Bogota	.06870445	.0000883	.0429792	-.0110338	.1581266	(P)
Cali	.06183682	-.0016488	.06200577	-.0617502	.1853082	(P)
Medellin	.04297218	-.0035567	.05462709	-.0727852	.1473747	(P)
Other	.04371865	-.0025466	.04333903	-.0458658	.1258189	(P)
IPW_X						
Bogota	.08359563	-.0039804	.05191273	-.025494	.1794194	(P)
Cali	.03992928	-.0000489	.07298156	-.0965727	.1782761	(P)
Medellin	.00830797	-.0144622	.0761305	-.1511961	.1418852	(P)
Other	.03200969	.0096506	.06160901	-.0716332	.1609599	(P)
Eqn1						
PM_X_NOV	.01811818	.0008978	.07551621	-.1097517	.1839608	(P)
IPW_X	.00791959	-.0096996	.09089705	-.1881287	.1795231	(P)

(P) percentile confidence interval

Caso 4: Caso 3 pero estimaciones especificas a cada tratamiento

- No es posible agregar efectos fijos a nivel de cursos para cada tratamiento porque los identifican perfectamente.

- El programa permite realizar las estimaciones para cada tratamiento por separado con variables comunes y especificas a cada uno de ellos.

- Incluimos el reporte de los rmsd, una medida de que tan distintos son los tratamientos entre si.

Definimos las mismas macros y variables que en el caso 3, salvo por las “xvar”

Caso 4: Caso 3 pero estimaciones especificas a cada tratamiento

Después de identificar los cursos por ciudad, definimos una macro con la especificación específica para cada ciudad

```
*Macro con las especificaciones para cada ciudad
local spec_bogota   "`bl_cov' i.curso_bogota"
local spec_cali     "`bl_cov' i.curso_cali"
local spec_medellin "`bl_cov' i.curso_medellin"
local spec_otro     "`bl_cov' i.curso_otro"
```

En el orden en que toman los valores de los tratamientos, se listan las especificaciones, separadas por “punto y coma” “**;**”

```
*Macro que agrupa las especificaciones de todas las ciudades
local specifications "`spec_bogota'; `spec_cali'; `spec_medellin'; `spec_otro'"
```

Caso 4: Caso 3 pero estimaciones especificas a cada tratamiento

Llamamos al comando

```
set seed 123456
mteffects `yvar' , tvar(`tvar') xvar(`xvar') xgps(`xgps')                ///
            estimators(`estimators')                                     ///
            treatment_lbls(`treatment_lbls')                             ///
            tvardif(`tvardif')                                           ///
            specifications(`specifications') rmsd
```

Especificaciones particulares a cada tratamiento

Medida de “distancia” entre tratamientos

Caso 4: Caso 3 pero estimaciones especificas a cada tratamiento

Outcome variable: **dformal_06**

	Observed Coef.	Bias	Bootstrap Std. Err.	[95% Conf.	Interval]	
PM_X_NOV						
Bogota	.06077873	-.0120484	.05734871	-.0592897	.1694314	(P)
Cali	.01549358	.0007278	.10591793	-.1954882	.226617	(P)
Medellin	.07458791	.0082632	.07160166	-.0626888	.2227007	(P)
Other	.03135076	-.0044358	.05427505	-.0703928	.1302327	(P)
IPW_X						
Bogota	.09269806	.0030141	.08094891	-.0751452	.2486473	(P)
Cali	.01166298	-.0120246	.13444911	-.2678589	.2384028	(P)
Medellin	.15219298	-.0178687	.11496724	-.1139126	.3502708	(P)
Other	.00160682	-.0001103	.0742837	-.1433907	.1405613	(P)
rmsd						
PM_X_NOV	.51250226	2.636857	4.9289104	.2472491	17.09262	(P)
IPW_X	.95614784	3.147936	8.0424209	.4354077	27.03845	(P)

(P) percentile confidence interval

Apéndice Caso 5: Comparar controles entre ciudades

Veamos que tan similares son los controles entre las distintas ciudades, y las características de los estos individuos comparables entre si

El comando nos permite entre otras cosas

- Observar el funcionamiento del balanceo de los dintintos tratamientos previo y posterior a la definición de la región de overlap.
- Generar histogramas de calidad de las regiones de overlap.
- Graficar las densidades de kernel de las distribuciones asociadas a cada grupo

Apéndice Caso 5: Comparar controles entre ciudades

- Al igual que en los casos anteriores, definimos las variables para las generación del gps y nuestra variable dependiente (que no va a ser utilizada)
- Eliminaremos los tratados de la muestra.
- Nos quedaremos únicamente con individuos de Bogotá y de Otras ciudades (No Cali ni Medellín)
- No reportaremos en este caso tablas de resultados.
- Características de las estimaciones (y los estimadores) no son necesarios.

Apéndice Caso 5: Comparar controles entre ciudades

Definimos un directorio donde guardar los resultados.

```
*Macro con el directorio donde se guardaran los resultados
local results "C:\Users\jmejalenko\Desktop\IEDW\Results"
```

Definimos un set de variables sobre las cuales calcularemos el balanceo.

```
*Macro con las bariables sobre las que analizamos el balanceo
local balvar "pempl_04 profit_04 tenure_04"
```

Definimos directamente en el comando algunas opciones

```
set seed 123456
mteffects `yvar' , tvar(`tvar') xgps(`xgps')
          treatment_lbls(`treatment_lbls')
          resultsdir(`results')
          percentile(Calidad_overlap)
          kernel(Kernel)
          balvar(`balvar')
          nobs
```

Directorio

Nombre p/ los histogramas

Nombre p/ las distribuciones

Variables a calcular el balanceo

Indica al programa que no es necesario realizar el bootstrap

Apéndice Caso 5: Balanceo

Medias de las variables seleccionadas

	raw_1	raw_2	raw_ovlp_1	raw_ovlp_2	IPW_1	IPW_2
pempl_04	0.89	0.81	0.89	0.80	0.85	0.85
profit_04	17278.35	31645.50	17278.35	33413.41	25237.98	25125.05
tenure_04	6.16	6.24	6.16	6.34	6.39	6.58

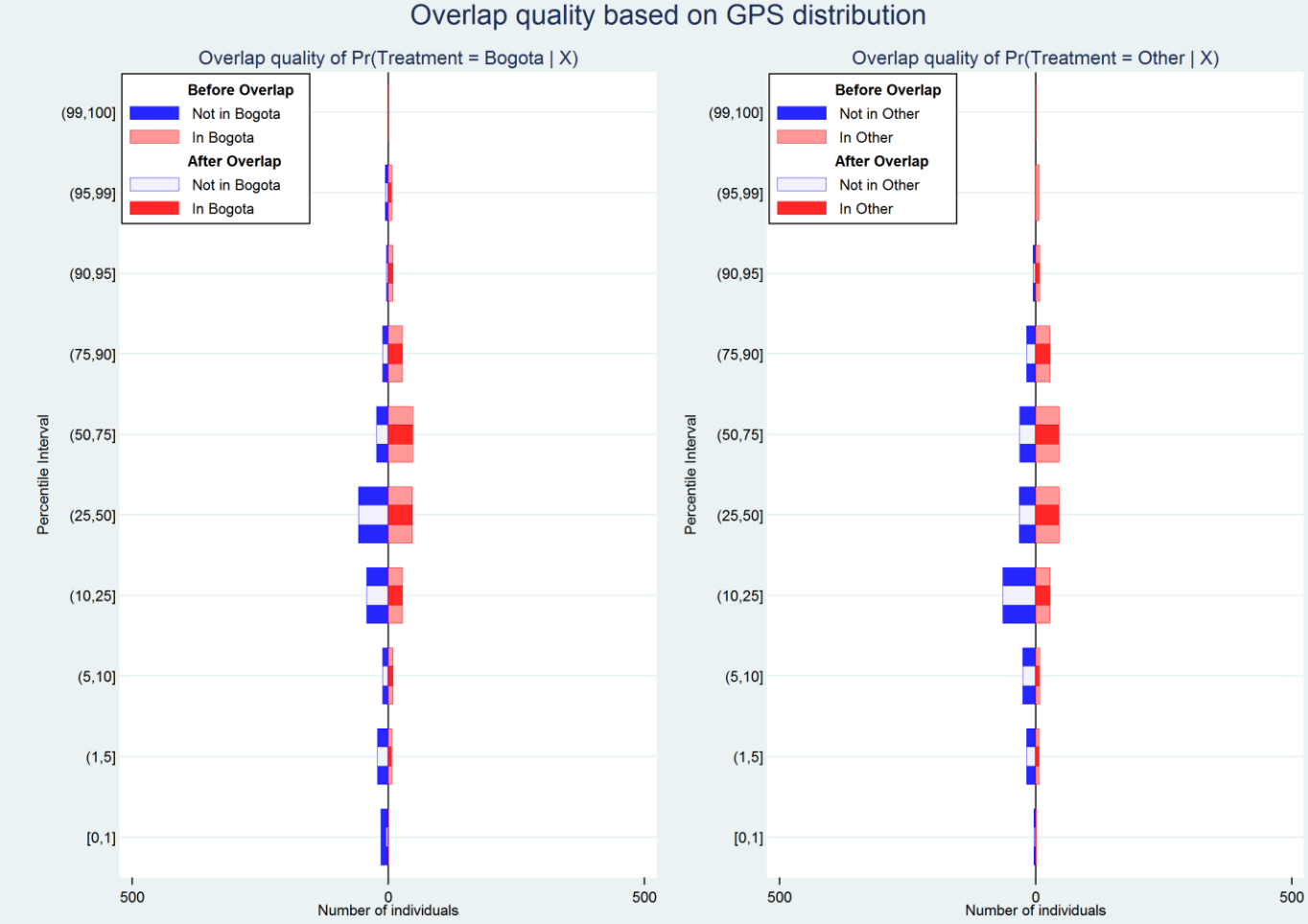
P-values

	Raw	Raw w/Ovlp	IPW
pempl_04	0.034	0.019	0.993
profit_04	0.064	0.044	0.989
tenure_04	0.911	0.812	0.814
nr_Xs_pval_lt_10_perc	2	2	0
nr_Xs_pval_lt_5_perc	1	2	0
nr_Xs_pval_lt_1_perc	0	0	0

RMSD

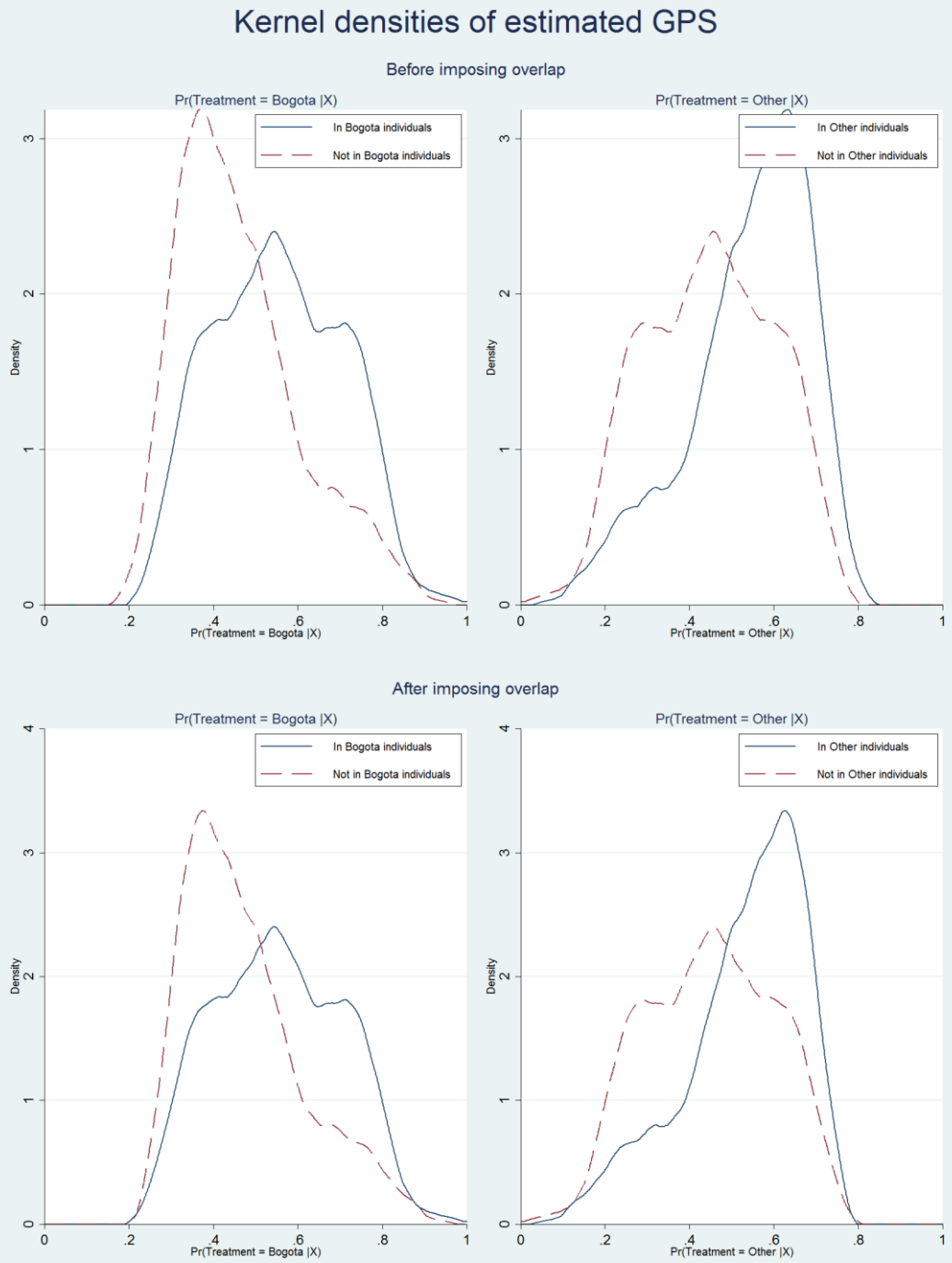
	Raw	Raw w/Ovlp	IPW
pempl_04	0.05	0.05	0.00
profit_04	0.29	0.32	0.00
tenure_04	0.01	0.01	0.01

Apéndice Caso 5:
Calidad de las regiones de overlap



Notes:
Outside wide bars and inside thin bars display the number of individuals whose GPS fall in a percentile interval, before and after imposing overlap, respectively.
Percentile intervals in each site are constructed using the GPS distribution, before imposing overlap, for individuals in that site only.

Apéndice Caso 5: Kernel densities



Summary

- We have seen several examples of extensions of the binary methods to the cases of continuous and multi-valued treatments
- Under randomization things are really not more complicated
- For non-experimental and quasi-experimental methods the econometrics is a little bit more complicated
- But, basically we do the same, deal with the issue of selection bias
- Apply the same sort of procedures: check for overlap, balancing, etc.

References

- Behrman, Jere R., Yingmei Cheng, and Petra E. Todd. 2004. “Evaluating Preschool Programs When Length of Exposure to the Program Varies: A Nonparametric Approach.” *Review of Economics and Statistics* 86 (1): 108–32.
- Bia, Michela, Carlos A. Flores, Alfonso Flores-Lagunes, and Alessandra Mattei. "A Stata package for the application of semiparametric estimators of dose–response functions." *Stata Journal* 14, no. 3 (2014): 580-604.
- Cattaneo, Matias D. 2010. “Efficient Semiparametric Estimation of Multi-Valued Treatment Effects under Ignorability.” *Journal of Econometrics* 155 (2): 138–54.
- Flores, Carlos A., Alfonso Flores-Lagunes, Arturo Gonzalez, and Todd C. Neumann. 2011. “Estimating the Effects of Length of Exposure to Instruction in a Training Program: The Case of Job Corps.” *Review of Economics and Statistics* 94 (1): 153–71.
- Flores, Carlos A., and Oscar A. Mitnik. 2013. “Comparing Treatments across Labor Markets: An Assessment of Nonexperimental Multiple-Treatment Strategies.” *Review of Economics and Statistics* 95 (5): 1691–1707.
- Hirano, Keisuke, and Guido W. Imbens. 2004. “The Propensity Score with Continuous Treatments.” In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, edited by Andrew Gelman and Xiao-Li Meng, 73–84. Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley and Sons.
- Imbens, Guido W. 2000. “The Role of the Propensity Score in Estimating Dose-Response Functions.” *Biometrika* 87 (3): 706–10.
- Imbens, Guido W. 2015. “Matching Methods in Practice.” Journal of Human Resources 50 (2) Spring: 373–419.

omitnik@iadb.org
jmejalenko@iadb.org

#devthatworks