Evaluación de Impacto: Synthetic Control

Francesco Bogliacino

Tabla de contenido

Origin of SBTC

- Exogenous (technology push argument)
  - ICT capital
  - Technology revolution;
  - General Purpose Technology introduction;
- Endogenous (demand pull argument)
  - Profit incentives of machine inventors
  - Skill complementing technologies

Industrial and Corporate Change, 2016, Vol. 25, No. 2, 227–243
doi: 10.1093/icc/dtv010
Advance Access Publication Date: 14 April 2015
Original Article

OXFORD

Endogenous skill biased technical change: testing for demand pull effect

Francesco Bogliacino[1,*] and Matteo Lucchese[2]

Natural Experiment

- High Share of tertiary education (and engineering/natural science bias) in East Germany;
- Control Sample: comparable countries;
- West Germany only

Test

SBTC shock

Initial shares in abstract and manual jobs
Abstract + Routine + Manual = 100%

$$\Delta w_{sejk\tau} = \sum_t \alpha_t^A \gamma_{sejk}^A \cdot 1[\tau = t] \cdot D_{it} + \sum_t \alpha_t^M \gamma_{sejk}^M \cdot 1[\tau = t] \cdot D_i +$$

$$\sum_t \beta_t^A \gamma_{sejk}^A \cdot 1[\tau = t] + \sum_t \beta_t^M \gamma_{sejk}^M \cdot 1[\tau = t] + \Lambda + e_{sejk\tau}$$

s, Sex;
e, Educational level (three);
j, Region;
k, age group (four);
tau, LIS wave

Positive: AA's polarization (machine replacing / SBTC)

Dummies:
Age, time, education, regions

Identification Strategy

- Natural experiment
- High Share of tertiary education (and engineering/natural science bias) in East Germany;
- Control Sample: comparable countries;
- West Germany only

Key assumptions

- the pattern of comparative advantage of East German workers is not systematically different from their West German colleagues, for a given type of skill;
- the pattern of comparative advantage of East German workers before the unification should not be correlated with the subsequent wage dynamics in West Germany

DATA

- Luxembourg Income Studies (Germany, France, and Spain)
- ISCO classification (10 groups) allocated to A, R, M
- ISCED classification (3 groups)
- Gross wages (imputed taxes where necessary)

Treatment

- Tertiary Education in East Germany is clearly higher (36% M – 31% F against 27,8 and 27,1), y. 1994 (decided before)
- Pattern comparative advantage treated and control groups is not statistically different (t-test unequal variance)
- Balancing of other characteristics

table I WLS Stacked First-Difference Estimates of the Relationship Between Demographic Group Occupational Distributions in 1989 and Subsequent Annual Changes in Demographic Groups' Mean Log Wages by LIS wave, 1989 — 2000

	Males	Females
	(1)	(2)
Abstract Occupation Share		
1989 Share x 1989–1994 dummy	−0.021	−0.084
	(0.056)	(0.056)
1989 Share x 1994–2000 dummy	0.061	0.199
	(0.040)*	(0.075)***
1989 Share x 1989–1994 dummy x West Germany	0.020	0.036
	(0.044)	(0.100)
1989 Share x 1994–2000 dummy x West Germany	0.048	0.009
	(0.043)	(0.097)
Manual Occupation Share		
1989 Share x 1989–1994 dummy	−0.034	−0.268
	(0.011)***	(0.150)*
1989 Share x 1994–2000 dummy	0.428	0.411
	(0.079)***	(0.105)***
1989 Share x 1989–1994 dummy x West Germany	0.097	0.047
	(0.146)	(0.182)
1989 Share x 1994–2000 dummy x West Germany	0.061	−0.143
	(0.158)	(0.185)
Lambda	Yes	Yes
Obs	142	136
R-squared	0.51	0.38
F-test	7.41	4.01
(Pvalue)	(0.00)	(0.00)

Source: LIS data for West Germany 1989, West Germany 1994, West Germany 2000, France 1989, France 1994, France 2000, Spain 1990, Spain 1995, Spain 2000. Each column represents a WLS regression of average (annual) rate of change of annual wages by cohort and year, where cohorts are defined by sex, age (25-34, 35-44, 45-54, 55-64), region of residence (North and South), country of residence, and educational attainment (ISCED 1, 2 and 3). Weights are calculated as sum by cell of the sample weights.

Abstract, Routine (the omitted category) and Manual jobs are defined according to ISCO classification, homogeneous among the three countries: Abstract includes ISCO classes managers, professionals, technicians and associate professionals, and skilled agricultural workers; Manual includes services and sales workers, elementary occupations and armed forces occupations.

Lambda includes time, country, region, age groups, and educational attainment fixed effects.

Standard errors in parenthesis. One, two or three stars indicate significance at ten, five and one percent respectively.

Robustness checks

Table II Summing up of the estimated demand pull effect

	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment Effect										
abstract-first wave	0.015	-0.050	0.019	-0.035	0.017	-0.031	-0.014	0.014	-0.006	-0.017
	(0.029)	(0.038)	(0.026)	(0.034)	(0.028)	(0.035)	(0.048)	(0.113)	(0.052)	(0.117)
abstract-second wave	0.045	0.005	0.046	0.016	0.044	0.018	0.017	-0.025	0.027	-0.047
	(0.028)	(0.036)	(0.025)*	(0.032)	(0.026)*	(0.033)	(0.045)	(0.107)	(0.048)	(0.111)
manual-first wave	0.023	-0.140	0.057	-0.088	0.046	-0.080	0.011	0.004	0.030	-0.079
	(0.099)	(0.069)**	(0.089)	(0.060)	(0.091)	(0.059)	(0.154)	(0.193)	(0.163)	(0.206)
manual-second wave	0.098	-0.148	0.140	-0.100	0.130	0.094	-0.040	-0.203	-0.016	-0.279
	(0.106)	(0.070)**	(0.095)	(0.060)	(0.095)	(0.060)	(0.160)	(0.196)	(0.169)	(0.209)
1989 Shares x Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lambda	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spain included	No	No	No	No	No	No	Yes	Yes	Yes	Yes
ISCO 6and 10	Yes	Yes	No	No	No	No	No	No	Yes	Yes
Weights	Sample	Sample	Sample	Sample	Cohort	Cohort	Cohort	Cohort	Cohort	Cohort

Source: LIS data for West Germany 1989, West Germany 1994, West Germany 2000, France 1989, France 1994, France 2000, Spain 1990, Spain 1995, Spain 2000. Each column represents a WLS regression of average (annual) rate of change of annual wages by cohort and year, where cohorts are defined by sex, age (25-34, 35-44, 45-54, 55-64), region of residence (North and South), country of residence, and educational attainment (ISCED 1, 2 and 3). Weights are either calculated as sum by cell of the sample weights (labeled sample in the table) or as mean start and end-year share of employment of each demographic group for each wave. (cohort in the Table). The treatment is a dummy for West Germany interacted with 1989 occupational shares and time dummies (the regressions are identical to those in Table 1).

Abstract, Routine (the omitted category) and Manual jobs are defined according to ISCO classification, homogeneous among the three countries: Abstract includes ISCO classes managers, professionals, technicians and associate professionals, and skilled agricultural workers; Manual includes services and sales workers, elementary occupations and armed forces occupations. The ISCO 6-10 row indicates that "skilled agricultural" and "armed force" are included.

Lambda includes time, country, region, age groups, and educational attainment fixed effects.Standard errors in parenthesis. One, two or three stars indicate significance at ten, five and one percent respectively.

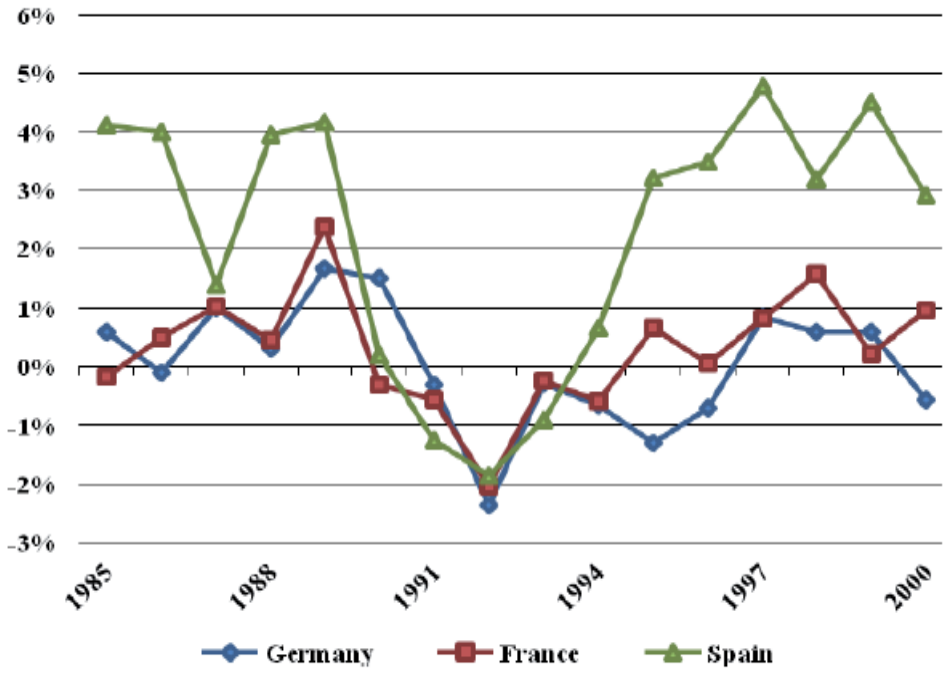
Confounding Factors

- Hours worked
- Labor Market Institutions
- Demand
- Taxation

Hours of work

- The ideal indicator would be hourly wages;
- Problem with reliability of self reported hours;
- Hours from related source show the same rate dynamics
- U shaped change is slightly more pronounced for Spain

Figure I The change in worked hours over the period 1989–2000 in the three countries

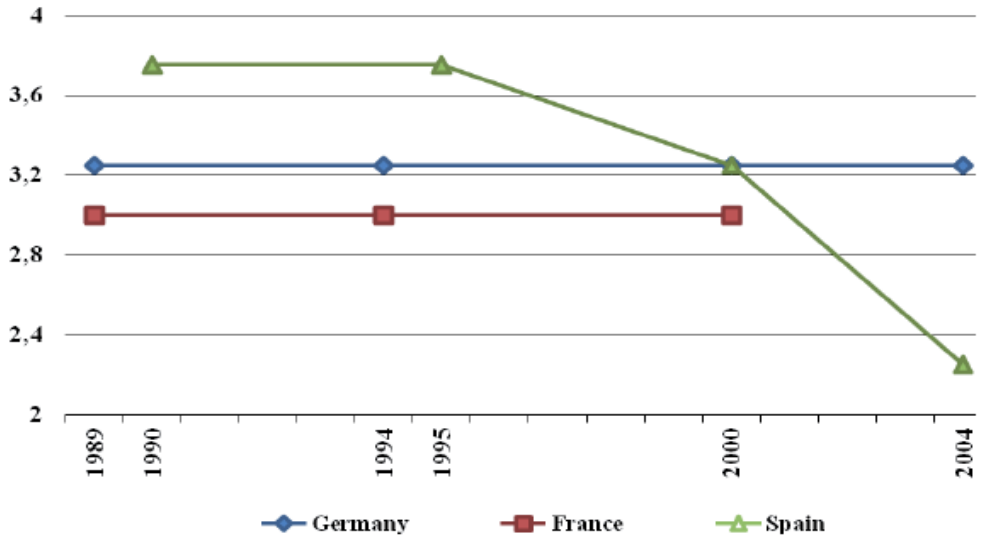


Source: EUKLEMS.

Flexibilisation affects the labour market dynamics

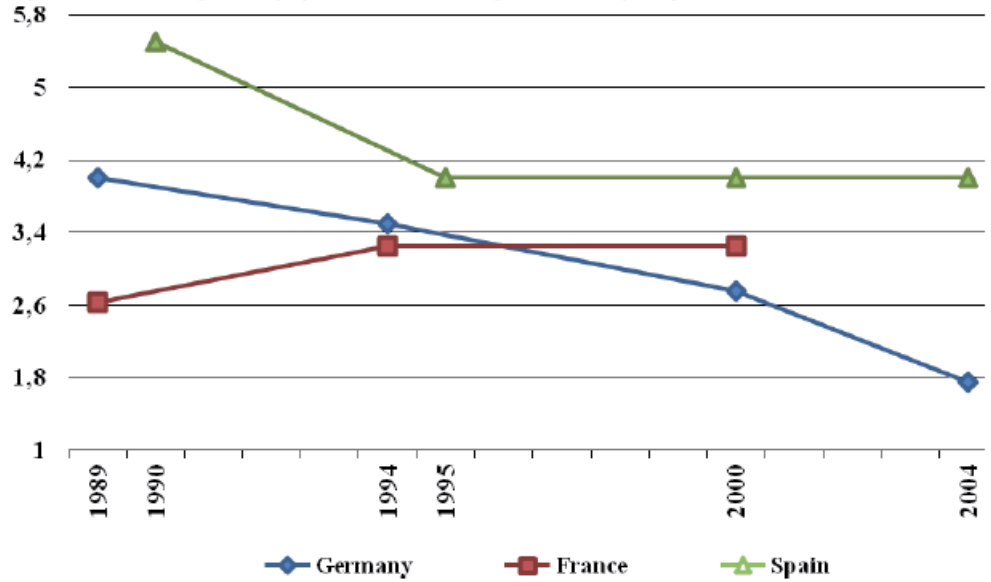
Need to control for it because Germany did introduce reforms while France did not

Figure II The change in Employment Protection Legislation. Regular contracts (EPR_v1 index)



Source: OECD EPL

Figure III The change in Employment Protection Legislation. Temporary contracts (EPT_v1)



Source: OECD EPL

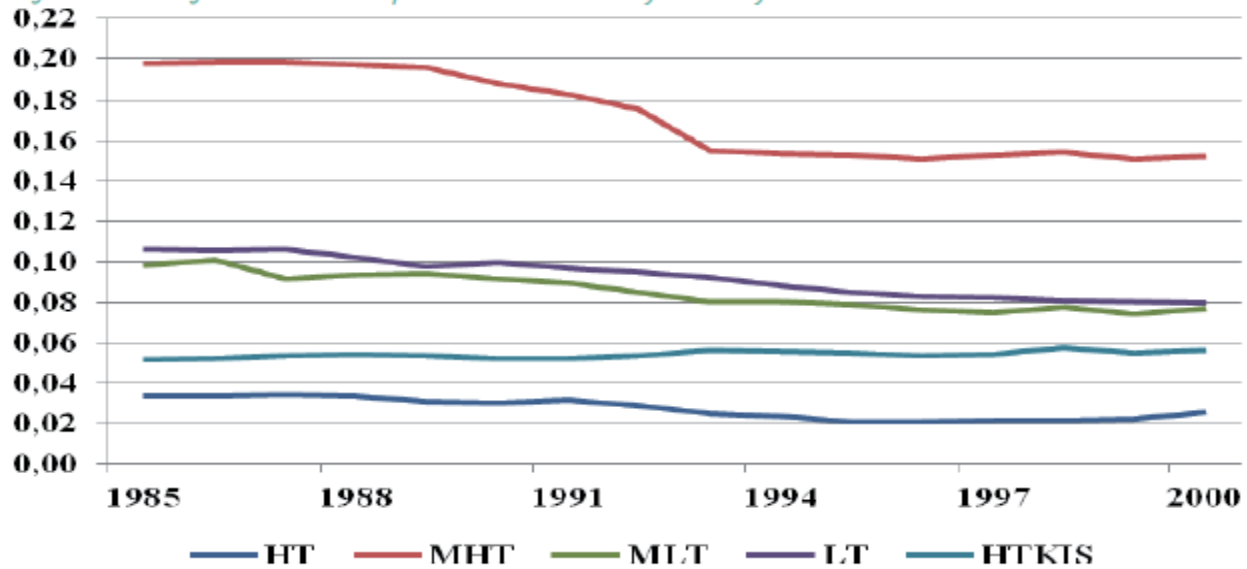
Demand

- Effective demand
- Sectoral composition (if strong diversity in tasks composition)

However

- Time and Country Fixed effects plus rate of change;
- Pattern of structural change is the same

Figure V Change in Sectoral Composition of the economy: Germany 1985-2000



Source: OECD STAN. The reported shares are computed in terms of value added with regard to the sum of manufacturing and market services. Sectors are defined according to NACE rev.3. High Tech (HT) is the sum of 30, 32, 33; Medium High Tech (MHT) is the sum of 24, 29, 31, 34, 35; Medium Low Tech (MLT) is the sum of 23, 25-28; Low Tech (LT) is the sum of 15-22 and 36-37; High Tech Knowledge Intensive Services (HTKIS) are 64, 72, 73.

Figure VI Change in Sectoral Composition of the economy: France 1985-2000

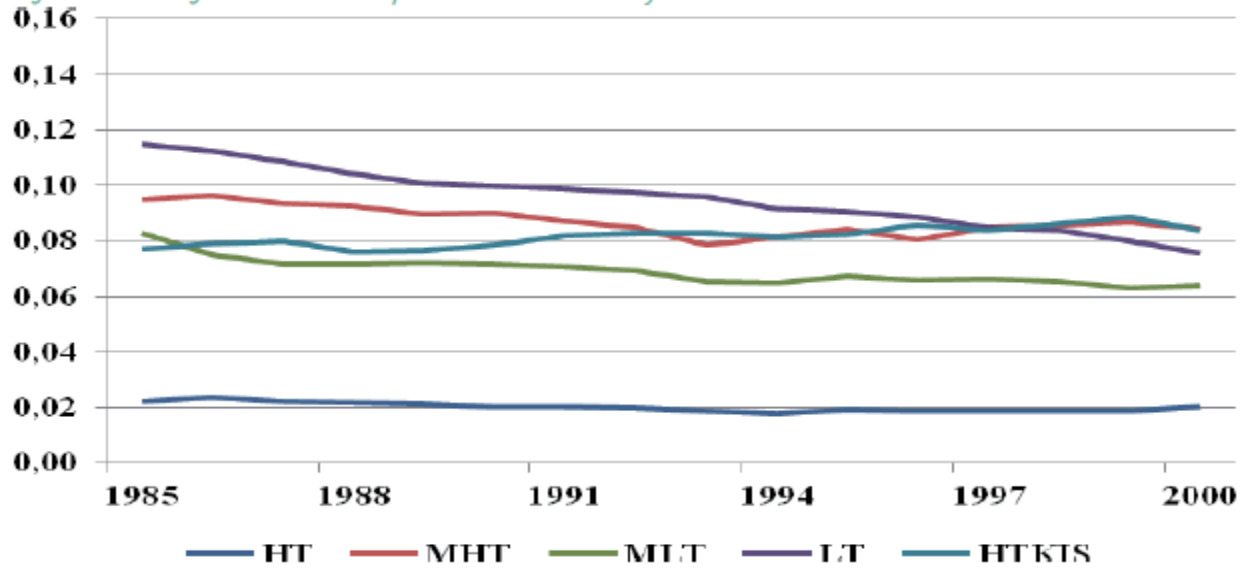
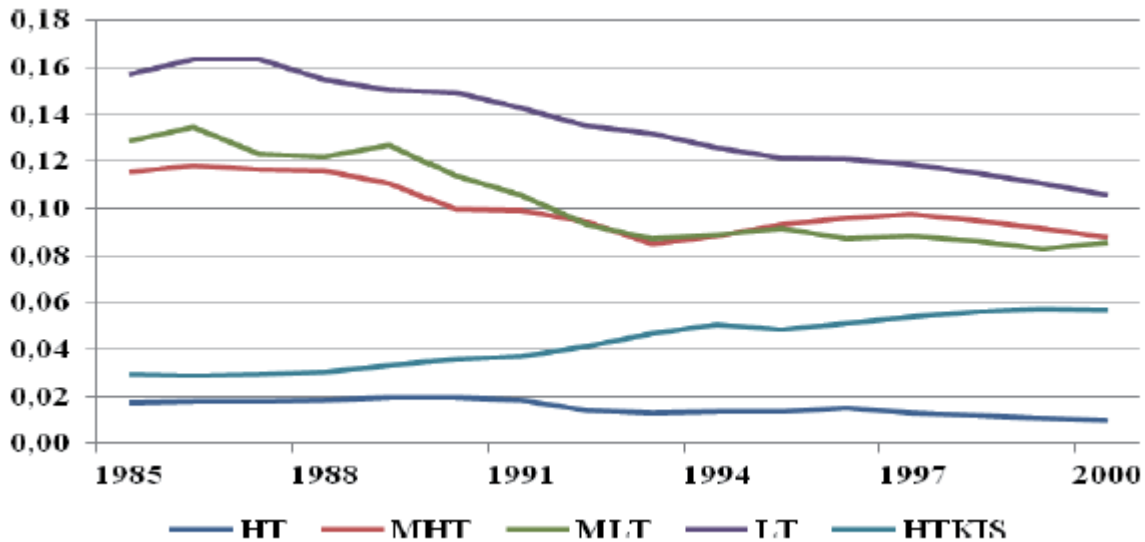


Figure VII Change in Sectoral Composition of the economy: Spain 1985-2000



Taxation

Figure IV The change in the marginal tax wedge between 1989 and 1994

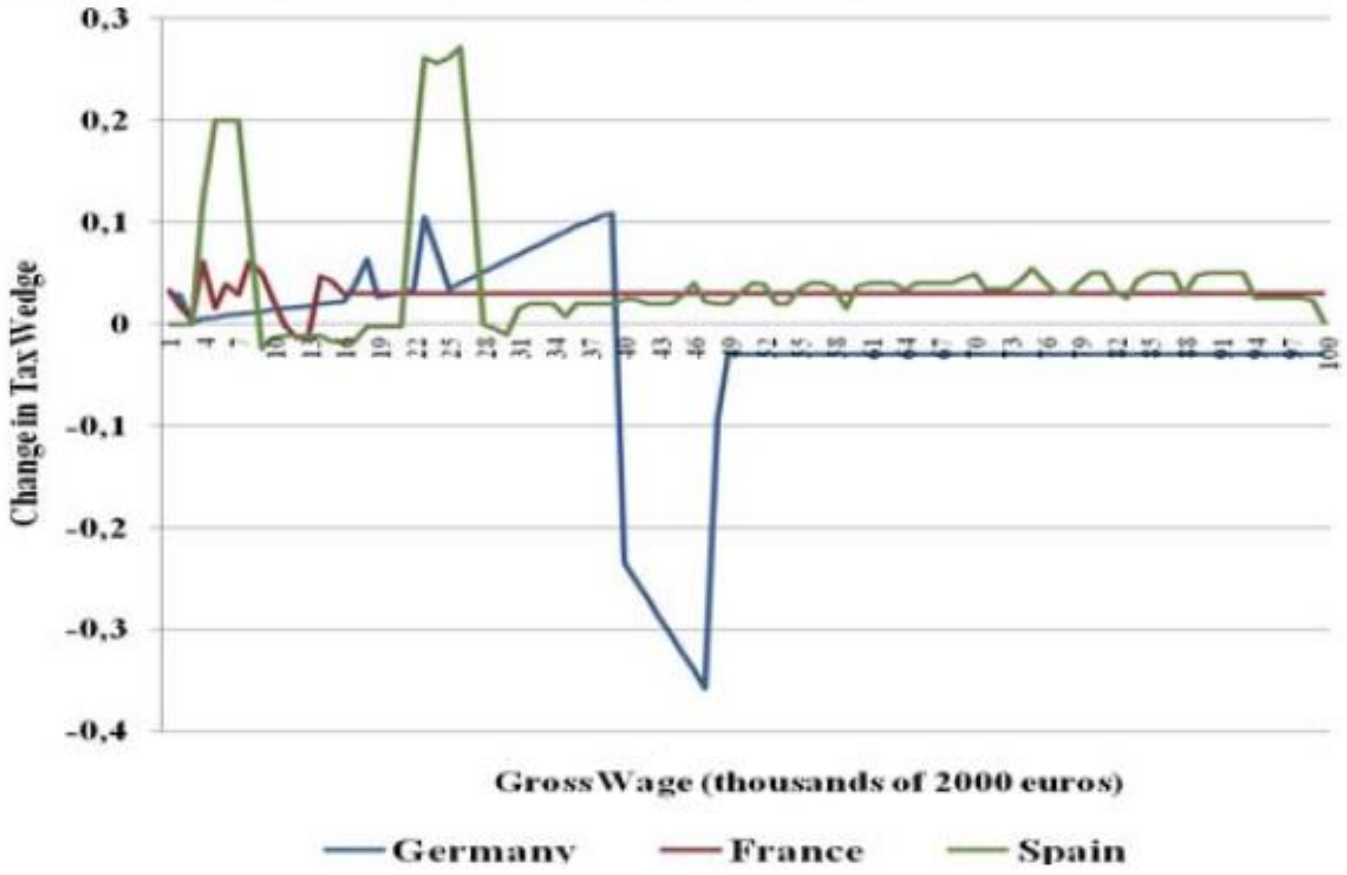


Tabla de contenido

Card 1990

Table 1. Characteristics of 16–61-Year-Olds in Miami, 1979.

Characteristic	*Whites*	*Blacks*	*Cubans*	*Hispanics*	*All*
Characteristics of Population Age 16–61					
1. Estimated Number (1000's)	319.3	244.1	252.4	102.9	928.4
2. Mean Education	12.8	11.4	11.0	11.6	11.8
3. Percent in Labor Force	75.6	68.3	77.2	68.8	73.1
Characteristics of Those in Labor Force					
4. Estimated Number (1000's)	241.3	166.6	194.7	70.8	678.2
5. Mean Education	13.1	11.8	11.3	11.9	12.1
6. Percent Age 16–24	21.1	24.1	22.0	26.0	22.8
Occupation Distribution (Percent of Employed)					
7. Professional and Technical	19.1	10.9	9.5	10.1	13.2
8. Managers	15.7	2.8	8.6	8.1	9.4
9. Sales	6.2	4.4	7.8	7.6	6.5
10. Clerical	21.9	21.0	19.1	20.9	20.9
11. Craftsmen	13.3	9.4	15.1	12.7	12.8
12. Operatives	4.4	8.4	19.4	16.7	11.1
13. Transportation Operatives	2.6	8.1	5.4	5.9	5.2
14. Laborers	5.1	10.5	4.7	4.0	6.3
15. Farm Workers	1.1	0.1	0.4	0.8	0.6
16. Less-Skilled Service Workers	5.0	13.3	6.1	10.2	8.0
17. More-Skilled Service Workers	5.7	10.9	4.0	3.0	6.2

Notes: White and black groups exclude hispanics. Hispanic group includes all hispanics other than Cubans. Less-skilled service workers include cleaning and food service workers. More-skilled service workers include health service, personal service, and protective service workers.

Source: Based on samples of employed workers in the outgoing rotation groups of the Current Population Survey in 1979.

Characteristic	*Mariel Immigrants*	*All other Cubans*
Educational Attainment (Percent of Population in Each Category):		
No High School	56.5	25.4
Some High School	9.1	13.3
Completed High School	9.5	33.4
Some College	6.8	12.0
Completed College	18.1	15.8
Percent Male	55.6	50.7
Percent Under 30 in 1980	38.7	29.6
Mean Age in 1980 (Years)	34.9	38.0
Percent in Miami in 1985	53.9	52.4
Percent Worked in 1984	60.6	73.4
Mean Log Hourly Earnings	1.37	1.71
Occupation Distribution (Percent Employed in Each Category):		
Professional/Managers	19.3	21.0
Technical	0.0	1.5
Sales	4.5	11.2
Clerical	2.5	13.5
Craftsmen	9.5	19.9
Operatives	19.1	13.8
Transportation Ops.	3.8	4.3
Laborers	10.8	3.3
Farm Workers	0.0	1.8
Less-Skilled Service	26.0	7.4
More-Skilled Service	4.6	2.3
Sample Size	50	528
Weighted Count	42,300	476,900

Table 3. Logarithms of Real Hourly Earnings of Workers Age 16–61 in Miami and Four Comparison Cities, 1979–85.

Group	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*
Miami:							
Whites	1.85	1.83	1.85	1.82	1.82	1.82	1.82
	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)	(.05)
Blacks	1.59	1.55	1.61	1.48	1.48	1.57	1.60
	(.03)	(.02)	(.03)	(.03)	(.03)	(.03)	(.04)
Cubans	1.58	1.54	1.51	1.49	1.49	1.53	1.49
	(.02)	(.02)	(.02)	(.02)	(.02)	(.03)	(.04)
Hispanics	1.52	1.54	1.54	1.53	1.48	1.59	1.54
	(.04)	(.04)	(.05)	(.05)	(.04)	(.04)	(.06)
Comparison Cities:							
Whites	1.93	1.90	1.91	1.91	1.90	1.91	1.92
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Blacks	1.74	1.70	1.72	1.71	1.69	1.67	1.65
	(.01)	(.02)	(.02)	(.01)	(.02)	(.02)	(.03)
Hispanics	1.65	1.63	1.61	1.61	1.58	1.60	1.58
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)

Note: Entries represent means of log hourly earnings (deflated by the Consumer Price Index—1980 = 100) for workers age 16–61 in Miami and four comparison cities: Atlanta, Houston, Los Angeles, and Tampa–St. Petersburg. See note to Table 1 for definitions of groups.

Source: Based on samples of employed workers in the outgoing rotation groups of the Current Population Survey in 1979–85. Due to a change in SMSA coding procedures in 1985, the 1985 sample is based on individuals in outgoing rotation groups for January–June of 1985 only.

Table 4. Unemployment Rates of Individuals Age 16–61 in Miami and Four Comparison Cities, 1979–85. (Standard Errors in Parentheses)

Group	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*
Miami:							
Whites	5.1	2.5	3.9	5.2	6.7	3.6	4.9
	(1.1)	(0.8)	(0.9)	(1.1)	(1.1)	(0.9)	(1.4)
Blacks	8.3	5.6	9.6	16.0	18.4	14.2	7.8
	(1.7)	(1.3)	(1.8)	(2.3)	(2.5)	(2.3)	(2.3)
Cubans	5.3	7.2	10.1	10.8	13.1	7.7	5.5
	(1.2)	(1.3)	(1.5)	(1.5)	(1.6)	(1.4)	(1.7)
Hispanics	6.5	7.7	11.8	9.1	7.5	12.1	3.7
	(2.3)	(2.2)	(3.0)	(2.5)	(2.1)	(2.4)	(1.9)
Comparison Cities:							
Whites	4.4	4.4	4.3	6.8	6.9	5.4	4.9
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.4)
Blacks	10.3	12.6	12.6	12.7	18.4	12.1	13.3
	(0.8)	(0.9)	(0.9)	(0.9)	(1.1)	(0.9)	(1.3)
Hispanics	6.3	8.7	8.3	12.1	11.8	9.8	9.3
	(0.6)	(0.6)	(0.6)	(0.7)	(0.7)	(0.6)	(0.8)

Note: Entries represent means of unemployment indicator variable for individuals age 16–61 in Miami and four comparison cities: Atlanta, Houston, Los Angeles, and Tampa–St. Petersburg. Samples are based on individuals in the labor force. See notes to Table 3 for definitions of groups and data sources.

Table 5. Means of Log Wages of Non-Cubans in Miami by Quartile of Predicted Wages, 1979–85. (Standard Errors in Parentheses)

Year	Mean of Log Wage by Quartile of Predicted Wage				*Difference of Means: 4th − 1st*
	1st Quart.	*2nd Quart.*	*3rd Quart.*	*4th Quart.*	
1979	1.31	1.61	1.71	2.15	.84
	(.03)	(.03)	(.03)	(.04)	(.05)
1980	1.31	1.52	1.74	2.09	.77
	(.03)	(.03)	(.03)	(.04)	(.05)
1981	1.40	1.57	1.79	2.06	.66
	(.03)	(.03)	(.03)	(.04)	(.05)
1982	1.24	1.57	1.77	2.04	.80
	(.03)	(.03)	(.03)	(.04)	(.05)
1983	1.27	1.53	1.76	2.11	.84
	(.03)	(.04)	(.03)	(.05)	(.06)
1984	1.33	1.59	1.80	2.12	.79
	(.03)	(.04)	(.04)	(.04)	(.05)
1985	1.27	1.57	1.81	2.14	.87
	(.04)	(.04)	(.04)	(.05)	(.06)

Note: Predicted wage is based on a linear prediction equation for the log wage fitted to individuals in four comparison cities; see text. The sample consists of non-Cubans (male and female, white, black, and Hispanic) between the ages of 16 and 61 with valid wage data in the earnings supplement of the Current Population Survey. Wages are deflated by the Consumer Price Index (1980 = 100).

Table 6. Comparison of Wages, Unemployment Rates, and Employment Rates for Blacks in Miami and Comparison Cities.
(Standard Errors in Parentheses)

	All Backs				*Low-Education Blacks*			
	Difference in Log Wages, Miami − Comparison		*Difference in Emp./Unemp., Miami − Comparison*		*Difference in Log Wages, Miami − Comparison*		*Difference in Emp./Unemp., Miami − Comparison*	
Year	*Actual*	*Adjusted*	*Emp. − Pop. Rate*	*Unemp. Rate*	*Actual*	*Adjusted*	*Emp. − Pop. Rate*	*Unemp. Rate*
1979	−.15	−.12	.00	−2.0	−.13	−.15	.03	−.8
	(.03)	(.03)	(.03)	(1.9)	(.05)	(.05)	(.04)	(3.8)
1980	−.16	−.12	.05	−7.1	−.07	−.07	.03	−8.2
	(.03)	(.03)	(.03)	(1.6)	(.05)	(.05)	(.04)	(3.5)
1981	−.11	−.10	.02	−3.0	−.05	−.11	.04	−7.7
	(.03)	(.03)	(.03)	(2.0)	(.05)	(.05)	(.04)	(4.2)
1982	−.24	−.20	−.06	3.3	−.17	−.20	−.04	.6
	(.03)	(.03)	(.03)	(2.4)	(.05)	(.05)	(.04)	(4.7)
1983	−.21	−.15	−.02	.1	−.13	−.11	.04	−3.3
	(.03)	(.03)	(.03)	(2.7)	(.06)	(.05)	(.04)	(4.7)
1984	−.10	−.05	−.04	2.1	−.04	−.03	.05	.1
	(.03)	(.03)	(.03)	(2.4)	(.06)	(.05)	(.04)	(4.7)
1985	−.05	−.01	−.06	−5.5	.18	.09	.00	−4.7
	(.04)	(.04)	(.04)	(2.6)	(.07)	(.07)	(.06)	(5.6)

Notes: Low-education blacks are those with less than 12 years of completed education. Adjusted differences in log wages between blacks in Miami and comparison cities are obtained from a linear regression model that includes education, potential experience, and other control variables; see text. Wages are deflated by the Consumer Price Index (1980=100). "Emp.-Pop. Rate" refers to the employment:population ratio. "Unemp. Rate" refers to the unemployment rate among those in the labor force.

Table 7. Means of Log Wages of Cubans in Miami: Actual and Predicted, and by Quartile of Predicted Wages. (Standard Errors in Parentheses)

	Mean of Log Wages Log in Miami			*Mean of Log Wages by Quartile of Predicted Wages*				*Mean Log Wage of Cubans Outside Miami*	*Difference in Cuban Wages, Miami − Rest-of-U.S.*	
Year	*Actual*	*Predicted*	*Actual-Predicted*	*1st*	*2nd*	*3rd*	*4th*		*Actual*	*Adjusted*
1979	1.58	1.73	−.15	1.31	1.44	1.64	1.90	1.71	−.13	−.10
	(.02)	(.02)	(.03)	(.02)	(.03)	(.04)	(.05)	(.04)	(.04)	(.04)
1980	1.54	1.68	−.14	1.25	1.49	1.59	1.81	1.66	−.12	−.06
	(.02)	(.02)	(.03)	(.02)	(.05)	(.04)	(.05)	(.03)	(.04)	(.03)
1981	1.51	1.68	−.17	1.23	1.43	1.55	1.80	1.63	−.13	−.09
	(.02)	(.02)	(.03)	(.03)	(.03)	(.04)	(.05)	(.03)	(.04)	(.03)
1982	1.49	1.68	−.19	1.27	1.43	1.50	1.77	1.71	−.22	−.12
	(.02)	(.02)	(.03)	(.03)	(.04)	(.04)	(.06)	(.03)	(.04)	(.03)
1983	1.48	1.65	−.17	1.16	1.41	1.56	1.80	1.62	−.14	−.08
	(.03)	(.02)	(.03)	(.02)	(.04)	(.04)	(.06)	(.03)	(.04)	(.03)
1984	1.53	1.69	−.17	1.20	1.40	1.65	1.88	1.63	−.10	−.08
	(.03)	(.02)	(.03)	(.03)	(.04)	(.05)	(.06)	(.03)	(.04)	(.03)
1985	1.49	1.67	−.18	1.19	1.43	1.53	1.80	1.77	−.27	−.19
	(.04)	(.03)	(.05)	(.06)	(.06)	(.08)	(.09)	(.06)	(.07)	(.05)

Notes: Predicted wage is based on a linear prediction equation for the log wage fitted to individuals in four comparison cities; see text. Predicted wages for Cubans in Miami are based on coefficients for Hispanics in comparison cities. The adjusted wage gap between Cubans in Miami and Cubans in the rest of the U.S. are obtained from a linear regression model that includes education, potential experience, and other control variables; see text. Wages are deflated by the Consumer Price Index (1980 = 100).

From CCT to SC

- El uso del studio de caso comparative es muy común (sobre todo en CP). Basado en el principio de las diferencias de Mill
- Es pensado por eventos agregados que queremos estudiar pero donde el enfoque estándar es limitado
- El asunto es como escoger las únidades de comparación y como hacer inferencia de la misma manera que se hace en estudios cuantitativos
- Existe un paralelo importante con las regresiones DiD. Al final siempre son ponderaciones entre unidades de control, pero en el synth los pesos son explicitadamente [0, 1]

La lógica

- Una unidad tratada, 1 (sin pérdida de generalidad), Y_{1t}, tratado desde T_0, pero observado en $0, 1, \dots, T_0, \dots, T_0 + T$
- J casos comparados (unidades desde 2 hasta J+1), no tratados, observados en $0, 1, \dots, T_0, \dots, T_0 + T$
- Para $T_0, \dots, T_0 + T$, dado un conjunto de pesos $w_i \in [0,1], \sum_i w_i = 1$, definimos $\sum_{i=2}^{J+1} w_i Y_{it}$ un control sintético

Synthetic control

- Abadie and Gardeazabal (2003) estiman el impacto de un ataque del ETA sobre crecimiento de Euskadi
- En vez de usar una unidad de comparación, se usan varias para acercarse más al comportamiento de la unidad tratada antes del tratamiento, tendremos en general k covariadas
- se trata de un control construido a partir de un conjunto de pesos. Los pesos nos dan la contribución de cada unidad en el conjunto de *donadores*
- no se hace extrapolación, como en las regresiones, porque se usa el *convex hull* de los donadores

Synthetic control

- Obviamente hay arbitrariedad en la elección del conjunto de donadores, en el modelo (X) y, sobretodo, del criterio de selección de los pesos

Tabla de contenido

Identificación

- Se trata de una mezcla entre DiD y control sobre observable
- Se define un conjunto $\mathbf{X}$ de covariadas para definer un espacio de cercanía entre unidades

Identification assumptions

- SUTVA
- CIA: una vez controlado por variables observables y trend especifico por unidad asignación es aleatoria
- Parallel trends: esto es impuesto haciendo emparejamiento sobre outcome pre-tratmiento

- Una unidad tratada, 1 (sin pérdida de generalidad), Y_{1t}, tratado desde T_0, pero observado en $0, 1, \dots, T_0, \dots, T_0 + T$
- J casos comparados (unidades desde 2 hasta J+1), no tratados, observados en $0, 1, \dots, T_0, \dots, T_0 + T$
- El ATT es

$$\delta_{1t} = Y_{1t}^{1} - Y_{1t}^{0}$$

Para cada $t = T_0, \dots, T_0 + T$

Optimal weights

- Define $W = \{w_2, \ldots, w_{J+1}\}', w_j \geq 0, \sum_j w_j = 1$
- Cada conjunto W es un control sintético
- Define X_1, X_0 dos matrices (kX1) y (kXJ) de las mismas variables para unidad tratada y donadores
- Tenemos que escoger los W^* que minimicen $\|X_1 - X_0 W\|$
- La estrategia de Abadie es definer una matriz V simétrica y semidefinida positive para minimizar

$$\|X_1 - X_0 W\| = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)} = \sum_{m=1}^{k} v_m (X_{1m} - \sum_{j=2}^{J+1} w_j X_{jm})^2$$

Optimal weights

- Obviamente V->W
- Se puede escoger basada en regresion (qué tanto predice), basada en argumento teórico, o en los datos (reducir los errores de predicción)

$$\sum_{t=1}^{T_0}\left(Y_{1t}-\sum_{j=2}^{J+1}W_j^*(V)Y_{jt}\right)^2$$

Si tenemos muchos años podemos entrenar $W_j^*(V)$ en un period y minimizar el MSPE en el periodo de validación

Tabla de contenido

Software

In Stata, the script can be installed by typing ssc install synth, replace, which downloads the software from the Statistical Software Components (SSC) archive. In R, the software is available as the Synth package from the Comprehensive R Archive Network at http://CRAN.R-project.org/package=Synth. The R package is also described in detail in Abadie, Diamond, and Hainmueller (2011). MATLAB code is available on the authors' websites. (Abadie et al. 2014, footnote 7)

Miremos en Stata

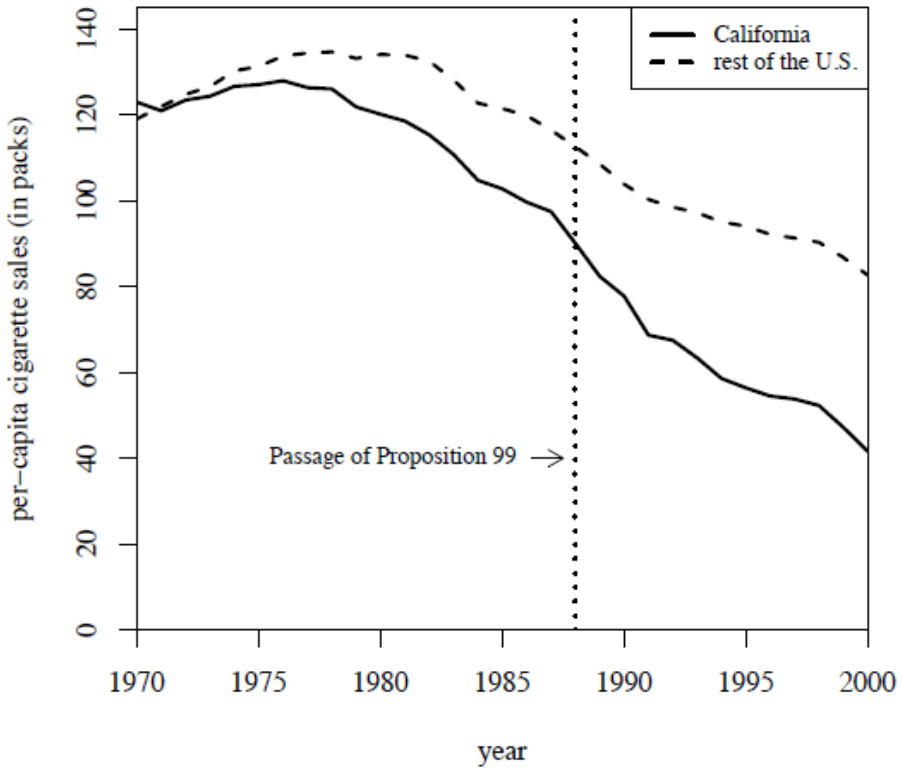
Les dejé el artículo para R en la Bibliografía en Dropbox

Tabla de contenido

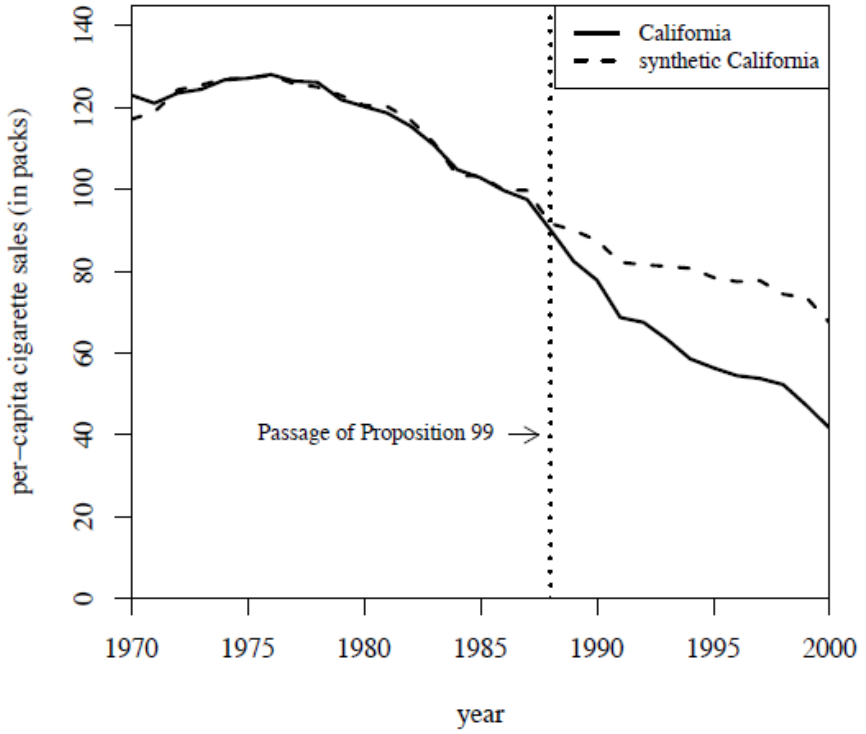
El ejemplo de Abadie et al. 2010

- California pasa la primera ley full anti tabaco en 1988
- Una serie de estados pasan leyes similares después (AK, AZ, FL, HI, MA, MD, MI, NJ, OR, WA, DC) y no pueden ser donadores

1) tendencias



2) Synth

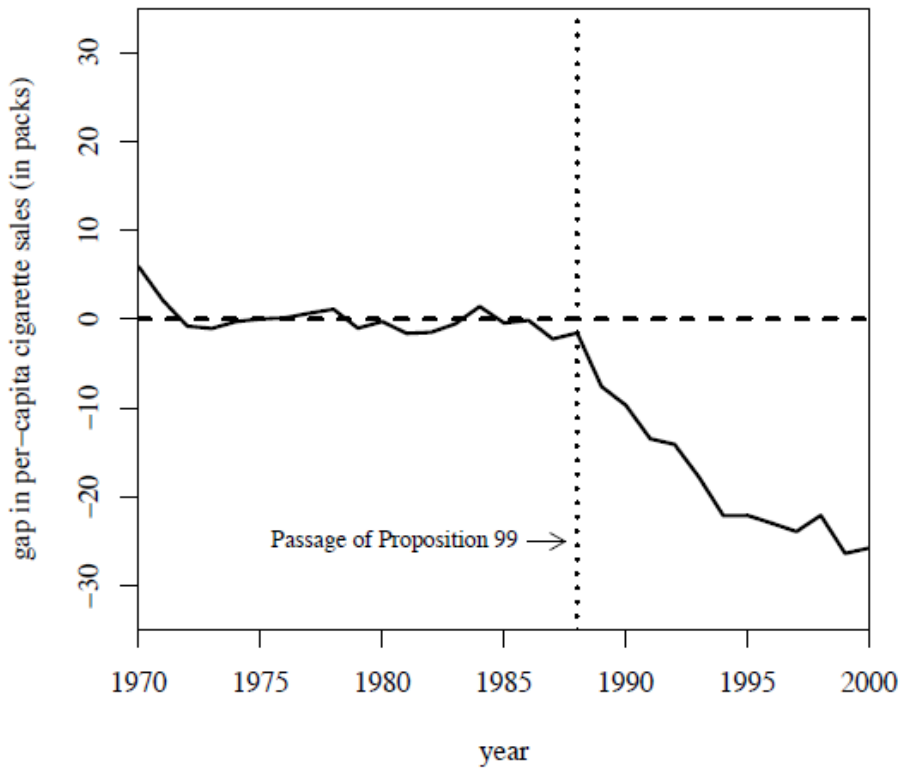

California
synthetic California
Passage of Proposition 99
per−capita cigarette sales (in packs)
0
20
40
60
80
100
120
140
1970
1975
1980
1985
1990
1995
2000
year

3) ajuste

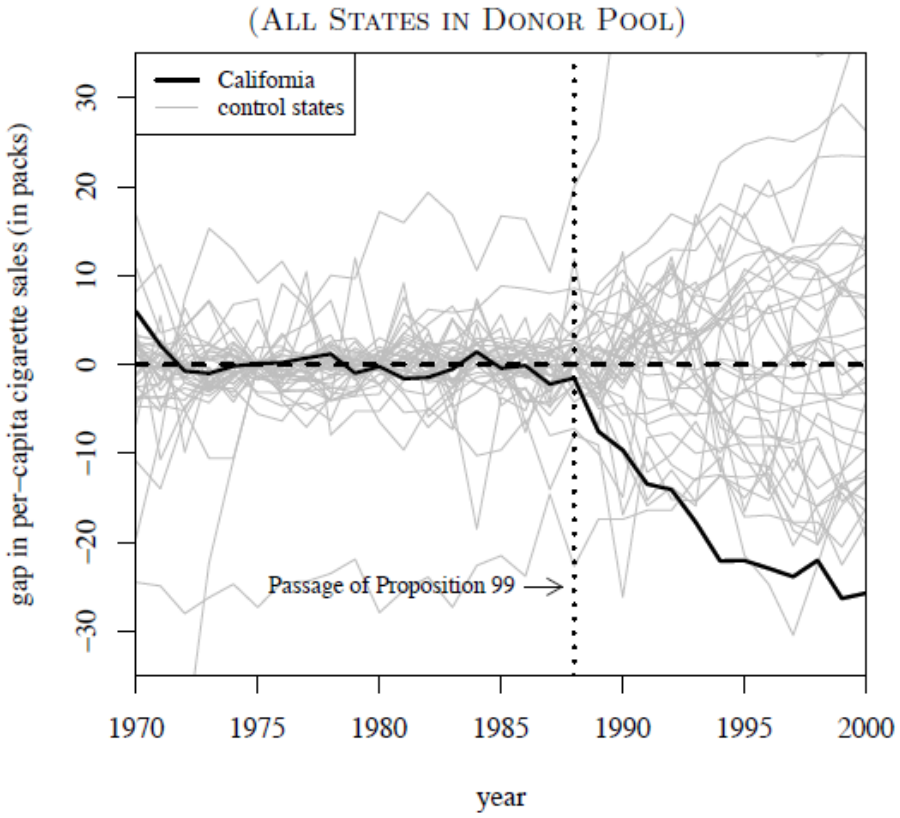
Variables	California		Average of
	Real	Synthetic	38 control states
Ln(GDP per capita)	10.08	9.86	9.86
Percent aged 15-24	17.40	17.40	17.29
Retail price	89.42	89.41	87.27
Beer consumption per capita	24.28	24.20	23.75
Cigarette sales per capita 1988	90.10	91.62	114.20
Cigarette sales per capita 1980	120.20	120.43	136.58
Cigarette sales per capita 1975	127.10	126.99	132.81

Note: All variables except lagged cigarette sales are averaged for the 1980-1988 period (beer consumption is averaged 1984-1988).

4) gap


gap in per−capita cigarette sales (in packs)
30
20
10
0
−10
−20
−30
Passage of Proposition 99
1970
1975
1980
1985
1990
1995
2000
year

5) Randomization inference



5b) randomization inference