Regional Policy Evaluation: Interactive Fixed Effects and Synthetic Controls

Laurent Gobillon
INED-Paris School of Economics

Thierry Magnac
Toulouse School of Economics

Online Appendix
March 11 2015

A Proofs and estimation method

A.1 Our implementation of Bai (2009) estimation method

Notations in this subsection slightly depart from the notations used in the text since for explaining Bai's method it is not necessary to distinguish treated and untreated observations or include a treatment indicator.

The interactive model can be rewritten in vector form at the individual level as:

$$y_i = x_i\beta + F'\lambda_i + \varepsilon_i, \tag{1}$$

where:

$$\begin{aligned} y_i &= (y_{i1}, ..., y_{iT})' \text{ of dimension } T \times 1, \\ x_i &= (x'_{i1}, ..., x'_{iT})' \text{ of dimension } T \times K, \\ \varepsilon_i &= (\varepsilon_{i1}, ..., \varepsilon_{iT})' \text{ of dimension } T \times 1, \\ F &= (f_1, ..., f_T) \text{ of dimension } L \times T. \end{aligned}$$

Model (1) can also be rewritten in matrix form:

$$Y = X \circledcirc \beta + F'\Lambda + \varepsilon, \tag{2}$$

where:

$$\begin{aligned} &Y = (y_1, ..., y_N) \text{ is a } T \times N \text{ matrix}, \\ &\Lambda = (\lambda_1, ..., \lambda_N) \text{ is a } L \times N \text{ matrix}, \\ &\varepsilon = (\varepsilon_1, ..., \varepsilon_N), \\ &X \text{ is a three-dimensional } T \times K \times N \text{ matrix}, \end{aligned}$$

and in which the sign $\circledcirc$ defines in an adhoc way the operation of X with β. Matrix $X \circledcirc \beta = [x_1\beta, ., x_N\beta]$ is of dimension $T \times N$.

Restrictions are needed to identify Λ and F since for any invertible $L \times L$ matrix Q, we can always write:

$$F'\Lambda = F'Q^{-1}Q\Lambda \tag{3}$$

and prove that (Λ, F) is observationally equivalent to $(Q\Lambda, Q^{-1\prime}F)$. Following the literature (e.g. Bai, 2009), we set:

$$\begin{array}{llll} FF'/T & = & I_L & L(L+1)/2 \text{ restrictions}, \\ \Lambda\Lambda' & \text{is} & \textit{diagonal} & L(L-1)/2 \text{ restrictions}. \end{array} \tag{4}$$

Parameters are estimated minimizing the least-square objective:

$$SSR\left(\beta,\Lambda,F\right)=\sum_{i=1}^{N}\left(y_i-x_i\beta-F'\lambda_i\right)'\left(y_i-x_i\beta-F'\lambda_i\right). \tag{5}$$

The minimization program can be solved using an iteration procedure. A minimizer in β can be computed given a value of F and Λ and minimizers in F and Λ can be computed given a value of β. It can also be shown that choosing initial values of F and Λ, and iterating leads to one solution which is the unique global minimizer. In line with Bai (2009), define the minimizer in β given parameters F and Λ as:

$$\beta\left(\Lambda,F\right)=\left(\sum_{i=1}^{N}x_i'x_i\right)^{-1}\left(\sum_{i=1}^{N}x_i'\left(y_i-F'\lambda_i\right)\right). \tag{6}$$

Note that the inverse of the matrix $\sum_{i=1}^{N}x_i'x_i$ can be computed once and for all, as it does not depend on the value of the parameters.

Conversely, given a value of β, values for F and Λ can be computed as follows. Let:

$$Z=Y-X\circledcirc\beta \text{ of dimension } T\times N. \tag{7}$$

The least square objective function (5) can be rewritten as:

$$Trace\left[\left(Z-F'\Lambda\right)'\left(Z-F'\Lambda\right)\right], \tag{8}$$

and the least-squares solution of Λ verifies:

$$\Lambda=\left(FF'\right)^{-1}FZ=FZ/T, \tag{9}$$

in which we have used the normalization $FF'/T=I_L$. Substituting (9) into (8) gives:

$$\begin{aligned}Trace\left[\left(Z-F'\Lambda\right)'\left(Z-F'\Lambda\right)\right] &= Trace\left[Z'(I-\frac{F'F}{T})(I-\frac{F'F}{T})Z\right]=Trace\left[Z'(I-\frac{F'F}{T})Z\right]\\ &= Trace\left(Z'Z\right)-Trace\left(Z'\frac{F'F}{T}Z\right)\\ &= Trace\left(Z'Z\right)-\frac{1}{T}Trace\left(FZZ'F'\right),\end{aligned} \tag{10}$$

in which we have used the invariance of the operator $Trace$ with respect to permutations of its arguments.

On the right-hand side, only the second term depends on F. Hence, an estimator of F is given by the maximization of $Trace\left(FZZ'F'\right)$. The estimator for F is the first L eigenvectors (multiplied by $\sqrt{T}$ because of the restriction $FF'/T = I_L$) associated with the L first largest values of the matrix:

$$ZZ' = \sum_{i=1}^{N} (y_i - x_i\beta)(y_i - x_i\beta)'.$$

Hence, given a value of β, F is obtained as a set of eigenvectors for a given matrix. Reciprocally, for a given value of F, β can be obtained from (6) after replacing λ_i using equation (9). The minimization program becomes:

$$\overline{SSR}(\beta, F) = \sum_{i=1}^{N} (y_i - x_i\beta)'(I - \frac{F'F}{T})(y_i - x_i\beta),$$

and the least-squares estimator of β is given by:

$$\beta(F) = \left(\sum_{i=1}^{N} x_i'(I - \frac{F'F}{T})x_i\right)^{-1}\left(\sum_{i=1}^{N} x_i'(I - \frac{F'F}{T})y_i\right).$$

Choosing some initial values for β or F, an iteration algorithm allows us to recover the OLS estimates of β and F. We finally estimate λ_i using equation (9).

A.2 Imposing constraints in estimation

We use an EM algorithm to obtain estimators of all parameters when using all observations of untreated units and pre-treatment observations of treated units only. Constraints on parameters can be imposed at further steps. The approach proceeds in the following way:

1. Step 0: Initialisation.

   We first apply Bai's estimation method developed above in the sample of the non treated group and we get preliminary and consistent estimates $\widehat{f}_t$, $\widehat{\Lambda}_U$, $\widehat{\beta}$. For the treated units before treatment, we perform the following OLS estimation:

   $$y_{it} - x_{it}\widehat{\beta} = \widehat{f}_t'\lambda_i + u_{it},$$

   and recover estimates of individual effects $\widehat{\lambda}_i$ for the treated group. We stack $\widehat{\lambda}_i$ into $\widehat{\Lambda}_T$ a consistent estimate of the corresponding matrix $\Lambda_T = (\lambda_1, ..., \Lambda_{N_1})$.

2. Step 1: Expectation Maximization Algorithm.

   - (E) We construct an expected value for the outcome in the treated group after treatment as if they were non treated (counterfactual):

     $$y_{it}(0) = x_{it}\widehat{\beta} + \widehat{f}_t'\widehat{\lambda}_i,$$

- (M) We reestimate Bai's model on the whole sample where the outcomes for the treatment group after treatment has been replaced by the (E) step. We thus recover consistent estimates $\widehat{f}_t^{(1)}$, $\widehat{\Lambda}_U^{(1)}$, $\widehat{\beta}^{(1)}$, $\widehat{\Lambda}_T^{(1)}$.

3. Step 2: Imposing constraints.

   Suppose that the set of weights derived for the synthetic control is $\omega^{(i)}$.

   We write for the control and treated groups:

$$
\begin{aligned}
y_{it} &= x_{it}\beta + \widehat{f}_t'^{(1)}\lambda_i + \varepsilon_{it}^{(1)},\ t = 1,.,T \text{ if } i \text{ not treated,} \\
y_{it} &= x_{it}\beta + \widehat{f}_t'^{(1)}\Lambda_U\omega^{(i)} + \varepsilon_{it}^{(1)},\ t = 1,.,T_D - 1 \text{ if } i \text{ treated.}
\end{aligned}
$$

   Estimating these equations by OLS, we recover estimates of β and Λ_U that we denote $\widehat{\beta}^{(2)}$ and $\widehat{\Lambda}_U^{(2)}$. To improve efficiency of the estimated factors f_t, an additional Newton Raphson step can be:

   (a) Step 3: Estimate f_t in the equations for the control and treated groups that are:

$$
\begin{aligned}
y_{it} - x_{it}\widehat{\beta}^{(2)} &= f_t'\widehat{\Lambda}_U^{(2)} + \varepsilon_{it}^{(2)}, t = 1,.,T \text{ if } i \text{ not treated,} \\
y_{it} - x_{it}\widehat{\beta}^{(2)} &= f_t'\widehat{\Lambda}_U^{(2)}\omega^{(i)} + \varepsilon_{it}^{(2)}, t = 1,.,T_D - 1 \text{ if } i \text{ treated.}
\end{aligned}
$$

B Monte-Carlo simulations

In this appendix, we present additional results of Monte-Carlo simulations when the DGP differs from the benchmark which is the following. We consider a model with both additive and interactive effects. The equation includes a treatment indicator whose coefficient is equal to .3. We have $N = 143$, $N_0 = 13$, $T = 20$, $T_D = 8$. All additive and interactive factors are drawn in a uniform distribution on [0,1]. When considering different supports for treated and non-treated units, individual factors of treated units, whether additive or multiplicative, are incremented by .5 (overlapping support) or 1 (disjoint support). The error is drawn in a zero mean and unit variance normal distribution.
Table B.1 presents the results when the error is drawn in a $\left[-\sqrt{3}, \sqrt{3}\right]$ uniform distribution (the variance is equal to 1) instead of a normal one. Results are qualitatively unchanged. Results when using a smaller number of periods T and T_D are reported in Table B.2. Biases remain similar except for the method "Interactive effects, counterfactual" for which the bias is large (due

to the fact that the number of pre-treatment periods is small). As expected, standard errors are larger. In Table B.3, we report the results when using a larger number of observations N and N_0, and obtain lower standard errors but slightly larger biases for the methods "Interactive effects, constrained" and "Synthetic controls".

$[Insert\ Figures\ B.1,\ B.2 \text{ and } B.3]$

We also experimented with heterogeneous treatments. The treatment effect is assumed to be correlated with the second individual factor, the first factor loading, λ_{i1} being the standard linear fixed effect. The heterogenous treatment parameter is thus written as $\alpha_i = \alpha + \left(\lambda_{i2} - \frac{1}{N_1}\sum_{i=1}^{N_1} \lambda_{i2}\right)$. This specific form is retained to make sure that the average treatment on the treated is equal to α by construction as we have: $\frac{1}{N_1}\sum_{i=1}^{N_1} \alpha_i = \alpha$. This yields results close to those with homogenous treatments.

In a different vein, we also experimented with departures from iid errors. First, to generate results in Table B.4, we consider that idiosyncratic errors follow an AR(1) process $\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + \varsigma_{i,t}$ in which shocks $\varsigma_{i,t}$ are iid and drawn in a normal distribution with zero mean and unit variance. We fix $\varepsilon_{i,1} = 0$. We experiment with different values of ρ from .1 to .9. The value of ρ is reported in the table at the top of the column. We find that biases are similar across simulations but standard errors increase as ρ increases.

$[Insert\ Figure\ B.4]$

Results in Table B.5 are obtained when considering two mass points for the variance of errors. Variance can be large, $V_h = 1 + a$, or small, $V_l = 1 - a$ with equal probability. Observations are still independently distributed but not identically. The ratio between the two variance values is $r = V_h/V_l$ and we have $a = (1 - r)/(1 + r)$. We experiment with values of r running from 1.2 to 3 reported in Table B.5 at the top of the column. We find that this kind of heteroskedasticity has nearly no effect on biases. It slightly affects standard errors but not much.

$[Insert\ Figure\ B.5]$

We also compare the asymptotic variance given in Bai (2009), and averaged across simulations, with the empirical variance computed from the simulations. Table B.6 gives the results when the number of factors varies. When there is only one additive individual effect and one interactive effect, the two variances take similar values. The discrepancy between the two variances increases as the number of interactive effects increases. Yet, as the average asymptotic variance remains stable, it is the sample variance only which increases.

[*Insert Figure B.6*]

These last two experiments make it clear that the number of periods, $T = 20$, that we consider is sufficiently large to avoid the small sample biases that Ahn et al. (2001) describe.

C Data Appendix

C.1 The sample

We use the same data as in Gobillon et al. (2012). After eliminating the very few observations for which some socio-economic characteristics are missing, we are able to reconstruct 8,831,456 unemployment spells in the period of interest that runs from July 1993 to June 2003. We aggregate unemployment spell data by municipalities and half-years. The municipality rate of exit to a job for a given half-year is the ratio between the number of exits to a job during the half-year and the number of unemployed workers at the beginning of the half-year. The municipality rates of entry into unemployment and exit for unknown reasons are defined in the same way.

We restrict the sample by eliminating municipalities that are too small to be eligible for an enterprise zone and Paris districts whose population is much larger than that of any treated municipality. This restriction leaves us with a sample of 271 municipalities (258 controls and 13 treatments) having between 8,000 and 100,000 inhabitants. There are no noticeable differences between this restricted sample and the full sample except for population size. Roughly speaking, an average of 90,000 unemployed workers find a job each half-year and this corresponds to about 300 exits per half-year in each municipality.

C.2 Descriptive statistics

Figure C.1 describes the raw data and reports the evolution of exit rates to a job in the sample of treated municipalities and in three control groups: a sample composed by non-treated municipalities in a restricted population range and two subsamples of non-treated municipalities located respectively at a distance within 5 kilometers, and within a band of 5 to 10 kilometers around a treated municipality. For readability, we draw a vertical line at first half of 1997 (half-year 8) when the policy started to be implemented. The curves for the control groups are broadly decreasing and exhibit parallel trends throughout the period. The curve for the treatment group slightly differs from those for the control groups between half-years 1 and 12 (second half of 1993 to first half of 1999). In particular, the exit rate to a job remains flat in the treatment group between half-years 7 and 8 (second half-year of 1996 and first half-year of 1997) when the policy is implemented whereas it is decreasing in the control groups. The estimation of the treatment parameter that we report in this paper is a way of formalizing and testing that these diverging trends are statistically significant.

[*Insert Figure C*.1]

None of these differences appears in the evolution of exit rates to non-employment and the evolution of exit rates for unknown reasons.

C.3 Propensity score

Next, we estimate a Probit model of enterprise zone designation as a function of municipality control variables among which are measures of physical job accessibility, the municipal composition of the population in terms of nationality or education, the unemployment rate, the proportion of young adults, and household income (proxying for the fiscal potential). We also include in the specification the smallest distance to another municipality comprising an enterprise zone. This is to account for the willingness of authorities to scatter enterprise zones more or less evenly throughout the region.

The results of our benchmark weighted Probit in which the weight is the (square root of the) number of unemployed workers in the municipality appear in the first column of Table C.1.

[*Insert Table C*.1]

In line with the selection criteria, the larger the average household income in the municipality or the smaller the proportion of persons without a high school diploma, the less likely the municipality comprises an enterprise zone although the effect is hardly significant for the latter variable. The higher the proportion of individuals below 25 years of age or the larger the size of the population, the larger the probability that the municipality contains an enterprise zone. We also find that the larger the density of jobs attainable in less than 60 minutes by private vehicle, the less likely it is that the municipality will be endowed with an enterprise zone. This is consistent with the targeting of places with relatively lower job accessibility. The effect of the distance to the nearest enterprise zone is not significant.

Using the results in the first column, we predicted the propensity score for each municipality. Interestingly, it reveals that the supports of the predicted propensity scores in the treated and control groups differ quite markedly as shown in Table C.2.

$[Insert\ Table\ C.2]$

The smallest predicted probability in the treatment group is equal to 0.08%, a low score which is consistent with political tampering in designation. In order to satisfy the common support condition (Smith and Todd, 2005), we further restrict the control group to municipalities whose predicted propensity scores are larger than the value 0.04% (see Table C.3). This restriction shrinks the control group by a factor of 2 and it now includes 135 municipalities (instead of 258), which is about ten times the number of treated municipalities (13).

In Table C.3, we report averages of explanatory variables in the treatment and control groups to assess whether those groups are balanced.

$[Insert\ Table\ C.3]$

Since the treatment group is small, it seems difficult to report averages in stratified samples defined by the propensity score levels (Smith and Todd, 2005). We rather report them globally even if results are less easy to interpret. The covariates of interest seem to be balanced in the two subsamples except for two variables: the proportion of college graduates and household income. This explains the differences in the propensity score averages between the control and treatment group. Nevertheless, the coefficient of designated municipalities in linear regressions of those covariates on the propensity score and the designation indicator is not significant even at the 10% level which indicates that samples are approximately balanced.

C.4 Robustness checks

Table C.4 reports the estimated effect of the enterprise zone program on entries and exits when controlling for the estimated propensity score. In the interactive effect and difference-in-differences approaches, this is done by including among the regressors the propensity score interacted with a trend t/T to mimic the presence of the propensity score in levels in the first difference equation as in Gobillon et al. (2012). We also include the propensity score among variables used in the construction of synthetic controls. Results obtained using the method "Interactive effect, treatment dummy" are very close to those obtained in the baseline case except when studying entries. In that case, the treatment effect estimate is larger when the specification includes four or less factors (including an additive one). Treatment effects for the other outcomes – exits to a job and exits for unknown reason – when using synthetic controls are now close to zero. Results obtained with difference in differences are similar to those obtained previously. In summary, once again, synthetic controls estimates are more sensitive to the specification than factor model estimates and difference-in-differences estimates. Overall, results are in line with those reported in the article when there is no control for the propensity score.

[*Insert Table C.*4]

REFERENCES

Bai, J., 2009, "Panel Data Models With Interactive Fixed Effects", *Econometrica*, 77(4), 1229-1279.

Gobillon, L., T., Magnac and H. Selod, 2012, "Do unemployed workers benefit from enterprise zones? The French experience", *Journal of Public Economics*, 96(9-10):881-892.

Smith, J.A.. and P. Todd, 2005, "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators", *Journal of Econometrics*, 125, 305-353.

Table B.1: Monte-Carlo results, variation of support, uniform errors

Support difference	0	.5	1
Interactive effects,	**0.024**	**-0.040**	**-0.092**
counterfactual	*0.021*	*-0.043*	*-0.101*
	[0.179]	[0.184]	[0.237]
Interactive effects,	**0.022**	**-0.040**	**-0.082**
treatment dummy	*0.024*	*-0.037*	*-0.090*
	[0.162]	[0.168]	[0.273]
Interactive effects,	**0.022**	n.a.	n.a.
matching	*0.026*	n.a.	n.a.
	[0.161]	n.a.	n.a.
Interactive effects,	**0.000**	**0.406**	**0.721**
constrained	*0.002*	*0.409*	*0.714*
	[0.111]	[0.128]	[0.237]
Synthetic controls	**-0.011**	**0.638**	**1.490**
	-0.014	*0.638*	*1.490*
	[0.106]	[0.120]	[0.179]
Diff-in-diffs	**0.026**	**-0.051**	**-0.125**
	0.028	*-0.048*	*-0.130*
	[0.143]	[0.136]	[0.135]

Data generating process: number of observations: $(N_1, N) = (13, 143)$, number of periods: $(T_D, T) = (8, 20)$, number of individual effects (including an additive one): $L = 3$, treatment parameter: $\alpha = .3$, time and individual effects of the non treated drawn in a uniform distribution $[0, 1]$, individual effects of the treated drawn in a uniform distribution $[0 + s, 1 + s]$ with $s \in \{0, .5, 1\}$ reported at the top of column, errors drawn in a uniform distribution $\left[-\sqrt{3}, \sqrt{3}\right]$.
Notes: Estimation methods are detailed in Section 4.1. $S = 1000$ simulations are used. The average (resp. median) estimated bias is reported in bold (resp. italic). The empirical standard error is reported in brackets.
Results for "Interactive effects, matching" are not reported when $s \in \{.5, 1\}$ as, in some simulations, some treated and non treated observations might be completely separated. As a consequence, the logit model used to construct the propensity score is not identified.

Table B.2: Monte-Carlo results, variation of the number of periods

Number of periods	$T = 20, T_D = 8$	$T = 10, T_D = 4$
Interactive effects,	**0.022**	**-0.128**
counterfactual	*0.027*	*0.009*
	[0.175]	[19.80]
Interactive effects,	**0.025**	**0.022**
treatment dummy	*0.026*	*0.016*
	[0.153]	[0.256]
Interactive effects,	**0.016**	**0.021**
matching	*0.016*	*0.022*
	[0.158]	[0.264]
Interactive effects,	**-0.001**	**-0.009**
constrained	*0.001*	*-0.014*
	[0.110]	[0.142]
Synthetic controls	**-0.013**	**-0.020**
	-0.011	*-0.029*
	[0.108]	[0.132]
Diff-in-diffs	**0.028**	**0.023**
	0.025	*0.030*
	[0.137]	[0.202]

Data generating process: number of observations: $(N_1, N) = (13, 143)$, number of periods: $(T_D, T) \in \{(4, 10), (8, 20)\}$ with (T_D, T) reported at the top of column, number of individual effects (including an additive one): $L = 3$, treatment parameter: $\alpha = .3$, time and individual effects drawn in a uniform distribution $[0, 1]$, errors drawn in a normal distribution with mean 0 and variance 1.
Notes: Estimation methods are detailed in Section 4.1. $S = 1000$ simulations are used. The average (resp. median) estimated bias is reported in bold (resp. italic). The standard error is reported in brackets.

Table B.3: Monte-Carlo results, variation of the number of units

Number of individuals	$N = 143, N_0 = 13$	$N = 275, N_0 = 25$
Interactive effects,	**0.007**	**0.003**
counterfactual	*0.009*	*0.003*
	[0.174]	[0.126]
Interactive effects,	**0.010**	**0.002**
treatment dummy	*0.012*	*0.005*
	[0.155]	[0.107]
Interactive effects,	**0.003**	**0.006**
matching	*0.003*	*0.006*
	[0.157]	[0.117]
Interactive effects,	**-0.007**	**0.047**
constrained	*-0.011*	*0.045*
	[0.111]	[0.078]
Synthetic controls	**-0.016**	**0.069**
	-0.017	*0.069*
	[0.109]	[0.075]
Diff-in-diffs	**0.016**	**0.000**
	0.017	*0.000*
	[0.137]	[0.098]

Data generating process: number of observations: $(N_1, N) \in \{(13, 143), (25, 275)\}$ with (N_1, N) reported at the top of column, number of periods: $(T_D, T) = (8, 20)$, number of individual effects (including an additive one): $L = 3$, treatment parameter: $\alpha = .3$, time and individual effects drawn in a uniform distribution $[0, 1]$, errors drawn in a normal distribution with mean 0 and variance 1.
Notes: Estimation methods are detailed in Section 4.1. $S = 1000$ simulations are used. The average (resp. median) estimated bias is reported in bold (resp. italic). The standard error is reported in brackets.

Table B.4: Monte Carlo results: AR(1) errors

AR(1) autocorrelation	.1	.3	.5	.7	.9
Interactive effects,	**0.015**	**0.029**	**-0.026**	**-0.016**	**-0.039**
linear counterfactual	*0.020*	*0.022*	*-0.021*	*-0.044*	*-0.001*
	[0.235]	[0.436]	[0.660]	[0.838]	[1.530]
Interactive effects,	**0.013**	**0.023**	**0.020**	**0.027**	**0.050**
linear	*0.016*	*0.023*	*0.025*	*0.022*	*0.036*
	[0.188]	[0.301]	[0.395]	[0.437]	[0.477]
Interactive effects,	**0.012**	**0.011**	**-0.015**	**-0.02**	**-0.007**
matching	*0.015*	*0.017*	*-0.003*	*-0.009*	*0.006*
	[0.193]	[0.285]	[0.369]	[0.545]	[1.020]
Interactive effects,	**-0.003**	**-0.008**	**-0.021**	**-0.018**	**0.014**
constrained	*0.000*	*-0.007*	*-0.016*	*-0.022*	*0.006*
	[0.120]	[0.140]	[0.178]	[0.259]	[0.494]
Synthetic controls	**-0.012**	**-0.015**	**-0.023**	**-0.019**	**0.017**
	-0.01	*-0.01*	*-0.024*	*-0.02*	*0.005*
	[0.117]	[0.136]	[0.177]	[0.260]	[0.495]
First difference	**0.024**	**0.027**	**0.022**	**0.014**	**0.035**
	0.029	*0.025*	*0.023*	*0.026*	*0.035*
	[0.147]	[0.183]	[0.236]	[0.312]	[0.490]

Data generating process: number of observations: $(N_1, N) = (13, 143)$, number of periods: $(T_D, T) = (8, 20)$, number of individual effects (including an additive one): $L = 3$, treatment parameter: $\alpha = .3$, time and individual effects drawn in a uniform distribution $[0, 1]$, errors drawn in an AR(1) process whose autocorrelation is reported at the top of column and innovations are white noise (normally distributed with mean 0 and variance 1).

Notes: Estimation methods are detailed in Section 4.1. $S = 1000$ simulations are used. The average (resp. median) estimated bias is reported in bold (resp. italic). The standard error is reported in brackets.

Table B.5: Monte Carlo results: heteroskedastic errors

Variance ratio	1.2	1.5	2	2.5	3
Interactive effects,	**0.013**	**0.014**	**0.021**	**0.012**	**0.015**
linear counterfactual	*0.014*	*0.014*	*0.021*	*0.011*	*0.016*
	[0.170]	[0.158]	[0.140]	[0.122]	[0.123]
Interactive effects,	**0.016**	**0.015**	**0.016**	**0.011**	**0.016**
linear	*0.018*	*0.012*	*0.013*	*0.012*	*0.014*
	[0.148]	[0.137]	[0.126]	[0.109]	[0.106]
Interactive effects,	**0.010**	**0.013**	**0.015**	**0.009**	**0.010**
matching	*0.015*	*0.008*	*0.015*	*0.011*	*0.015*
	[0.152]	[0.144]	[0.128]	[0.116]	[0.114]
Interactive effects,	**-0.002**	**-0.001**	**0.000**	**-0.002**	**0.002**
constrained	*-0.003*	*-0.006*	*-0.002*	*-0.003*	*0.004*
	[0.107]	[0.100]	[0.093]	[0.081]	[0.078]
Synthetic controls	**-0.011**	**-0.009**	**-0.01**	**-0.009**	**-0.007**
	-0.014	*-0.01*	*-0.011*	*-0.008*	*-0.003*
	[0.106]	[0.098]	[0.092]	[0.081]	[0.077]
First difference	**0.021**	**0.020**	**0.019**	**0.016**	**0.021**
	0.018	*0.018*	*0.015*	*0.016*	*0.020*
	[0.132]	[0.121]	[0.111]	[0.097]	[0.093]

Data generating process: number of observations: $(N_1, N) = (13, 143)$, number of periods: $(T_D, T) = (8, 20)$, number of individual effects (including an additive one): $L = 3$, treatment parameter: $\alpha = .3$, time and individual effects drawn in a uniform distribution $[0, 1]$, errors drawn in a normal distribution with mean 0 and a two-point-of-support variance: $V_h = 1 + a$ and $V_l = 1 - a$. Observations are randomly allocated to large and small variance with equal probability 0.5. The variance ratio $r = V_h/V_l$ is reported at the top of the column.

Notes: Estimation methods are detailed in Section 4.1. $S = 1000$ simulations are used. The average (resp. median) estimated bias is reported in bold (resp. italic). The standard error is reported in brackets.

Table B.6: Monte-Carlo results: empirical variance and theoretically derived asymptotic variance

Number of individual effects	2	3	4	5	6
Mean bias	**0.017**	**0.011**	**0.010**	**0.019**	**0.022**
Empirical Standard error	*0.142*	*0.152*	*0.165*	*0.181*	*0.191*
Square root Asymp. Var	0.131	0.129	0.128	0.129	0.128

Data generating process: number of observations: $(N_1, N) = (13, 143)$, number of periods: $(T_D, T) = (8, 20)$, number of individual effects (including an additive one): $L \in \{2, 3, 4, 5, 6\}$ with L reported at the top of column, treatment parameter: $\alpha = .3$, time and individual effects drawn in a uniform distribution $[0, 1]$, errors drawn in a normal distribution with mean 0 and variance 1.

Notes: Mean bias: average bias computed across simulations; Empirical Standard error: empirical standard error computed across simulations; Square root Asymp. Var: square root of the average of the asymptotic variance computed across simulations.

Table C.1: Propensity score: effect of municipality characteristics on the probability of designation to receive an enterprise zone

	Weighted	Unweighted
Job density, 60 minutes by private vehicle	-3.999*	-4.171*
	(2.109)	(2.298)
Proportion of no diploma	37.779*	24.029
	(22.249)	(22.865)
Proportion of technical diplomas	20.998	0.974
	(28.215)	(28.900)
Proportion of college diplomas	38.978	17.299
	(29.889)	(31.336)
Distance to the nearest EZ	-0.027	-0.035
	(0.024)	(0.024)
Proportion of individuals below 25 in 1990	17.125***	11.834**
	(5.156)	(5.256)
Population in 1990	0.021**	0.019*
	(0.009)	(0.011)
Average net household income in 96	-4.975***	-2.033
	(1.563)	(1.593)
Constant	-32.115	-16.526
	(21.818)	(22.537)
Nb. observations	271	271
Pseudo-R^2	.542	.477

Note: ***: significant at 1% level; **: significant at 5% level; *: significant at 10% level. Probit estimation. The dependent variable is a dummy equal to one if the municipality is designated to receive an EZ (and zero otherwise). The sample is restricted to municipalities whose population is between 8,000 and 100,000 in 1990. The first column reports results when weighting by the square root of the number of unemployed workers at risk at the beginning of half-year 8 (1st half year of 1997), and the second column when using no weight.

Table C.2: Frequency of non-treated municipalities by propensity score brackets which bound are values of treated municipalities

Score bracket	Number of non-treated municipalities
[0, 0.0008)	125
[0.0008, 0.0119)	60
[0.0119, 0.1161)	52
[0.1161, 0.1772)	7
[0.1772, 0.3111)	6
[0.3111, 0.4404)	4
[0.4404, 0.4765)	0
[0.4765, 0.6091)	2
[0.6091, 0.7723)	2
[0.7723, 0.7933)	0
[0.7933, 0.8537)	0
[0.8537, 0.9032)	0
[0.9032, 0.9949)	0
[0.9949 , 1]	0
Total	258

Note: The observation unit is a municipality having between 8,000 and 100,000 inhabitants. The propensity score is computed as the predicted probability as derived from Probit estimates reported in Table C.1, column (1). Each bracket bound (excluding 0 and 1) corresponds to the propensity score of one treated municipality, with treated municipalities being sorted in ascending order of their estimated propensity score.

Table C.3: Average of municipality characteristics in treatment and control groups

	Treatment group	Control group, propensity score > 0.0004
Job density, 60 minutes by public transport	0.838	0.850
	(0.119)	(0.119)
Proportion of no diploma	0.536	0.465
	(0.041)	(0.074)
Proportion of technical diplomas	0.222	0.219
	(0.009)	(0.031)
Proportion of college diplomas	0.122	0.179
	(0.025)	(0.075)
Distance to the nearest EZ	9.074	11.016
	(12.193)	(8.051)
Proportion of individuals below 25 in 1990	0.416	0.372
	(0.038)	(0.043)
Population in 1990	45.201	43.578
	(18.226)	(26.357)
Average net household income in 96	0.375	0.509
	(0.087)	(0.125)
Number of observations	13	135

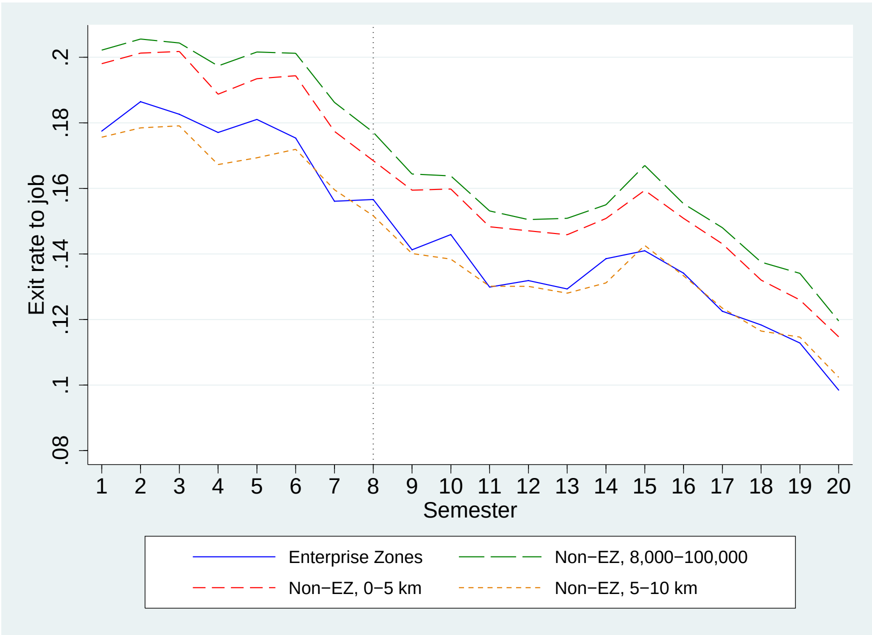
Note: Standard errors are reported in parenthesis under the means. The observation unit is a municipality between 8,000 and 100,000 inhabitants. Only municipalities whose estimated propensity score is greater than 0.0004 are considered in the control group. The propensity score is computed using the estimated coefficients of Table C.1, column (1).

Table C.4: Estimated enterprise zone program effects on unemployment exits and entry when controlling for the propensity score

Nb of indiv. effects	2	3	4	5	6
Exit rate to a job					
Interactive effects,	0.033	0.034	0.037	0.039	0.046
treatment dummy	[-0.004 ; 0.070]	[-0.005 ; 0.073]	[0.002 ; 0.072]	[0.004 ; 0.074]	[0.013 ; 0.079]
Synthetic controls			-0.006		
			[-0.048 ; 0.062]		
Diff-in-diffs			0.041		
			[0.008 ; 0.074]		
Exit rate for unknown reasons					
Interactive effects,	0.022	0.008	0.006	0.015	0.014
treatment dummy	[-0.017 ; 0.061]	[-0.027 ; 0.043]	[-0.027 ; 0.039]	[-0.018 ; 0.048]	[-0.017 ; 0.045]
Synthetic controls			0.004		
			[-0.066 ; 0.040]		
Diff-in-diffs			0.021		
			[-0.012 ; 0.054]		
Entry rate					
Interactive effects,	0.021	0.028	0.029	0.008	0.008
treatment dummy	[-0.010 ; 0.052]	[-0.001 ; 0.057]	[0 ; 0.058]	[-0.023 ; 0.039]	[-0.021 ; 0.037]
Synthetic controls			0.003		
			[-0.042 ; 0.022]		
Diff-in-diffs			0.027		
			[0.002 ; 0.052]		

Notes: Outcomes are computed in logarithms at the municipality level. The number of observations are $(N_1, N) = (13, 148)$ and the number of periods are $(T_D, T) = (8, 20)$. The estimated coefficient is the first reported figure. Its 95% confidence interval is given below in brackets. For the estimation method *Interactive effects, treatment dummy*, the confidence interval is computed considering that errors are independently and identically distributed. For the estimation method *Diff-in-diffs*, the feasible general least square estimator is computed assuming a constant within-municipality unrestricted covariance matrix. For *Synthetic controls*, the confidence interval is computed as explained in the text under the assumption of independently and identically distributed errors.

Figure C.1: Exit rate to a job, by group of municipalities



Note: The exit rates to a job are reported for half-years between 1 (2nd half year of 1993) and 20 (1st half year of 2003). Half-year 8 (1st half year of 1997) is the first half-year during which some municipalities are treated. Non-EZ: municipalities which do not include an EZ. 8,000-100,000: population between 8,000 and 100,000 in 1990. 0-5km: between 0 and 5km of a municipality including an EZ. Enterprise zones: municipalities which include an EZ.