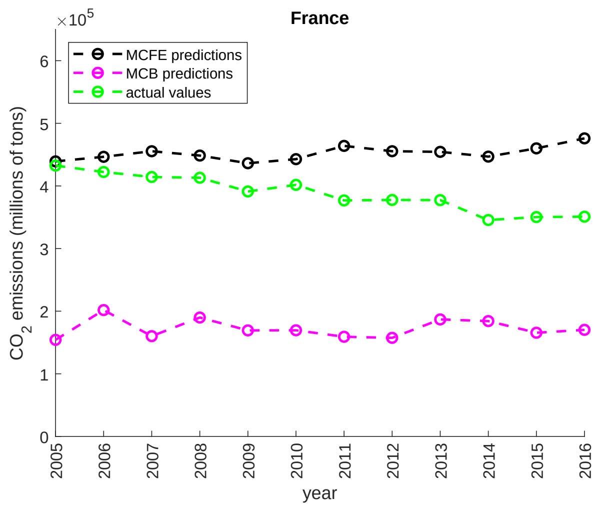
Supplementary material of

The Impact of the European Union Emissions Trading System on Carbon Dioxide Emissions: A Matrix Completion Analysis

Francesco Biancalani, Giorgio Gnecco, Rodolfo Metulini, Massimo Riccaboni

Supplementary Section A Comparison between Matrix Completion With/Without Fixed Effects Estimation

Supplementary Figure A.1 shows, through one representative example, that standard Matrix Completion (MC) without fixed effects estimation – i.e., the baseline MC (or MCB) method, which is obtained from the optimization problem (1) by further constraining the optimization vectors $\hat{\boldsymbol{\Gamma}}$ and $\hat{\boldsymbol{\Delta}}$ to be equal to $\mathbf{0}$ – systematically underestimated the missing elements of the matrix, when that method was applied to the specific data set considered in the present work. This did not occur, instead, when MC with Fixed Effects estimation (MCFE) was employed. In other words, of the two MC methods, only the latter turned out to be suitable for assessing the impact of the EU ETS policy (which was expected to be typically positive, at least after some years since the adoption of the policy).



Supplementary Figure A.1. For one MC repetition and the subsets of the test set associated with one representative treated countries: predictions obtained by MC with/without fixed effects estimation (i.e., by MCFE and MCB, respectively), and corresponding actual values. The reader is referred to Supplementary Section A for an additional discussion of this figure.

An additional comparison between MCFE and MCB was performed by[1], showing that MCFE achieved better reconstruction accuracy than MCB in the pre-treatment period (for which a ground truth is available). The work included also some sensitivity analyses (e.g., with respect to changing the size of the training set). It is also worth bringing to the reader's attention the recent study by[2], which investigates theoretically the impact of various parameters, such as data heterogeneity, size of the missing block, and deviation from ideal low rankness of the matrix to be reconstructed, on the accuracy of nuclear-norm based MC, for the case of data that are Missing Not At Random (MNAR). This is related to our study, being the MNAR assumption valid for our test set.

Supplementary Section B Comparison with the Difference-in-Differences (DiD) Model

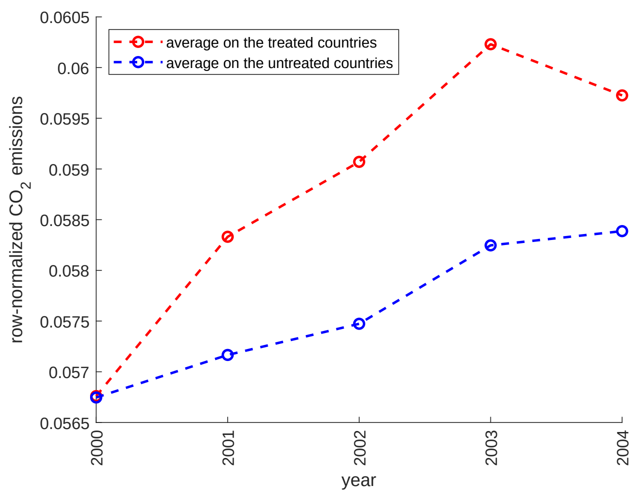
We also compared the MCFE results with those obtained by one traditional econometric method used to estimate counterfactuals, i.e., the Difference-in-Differences (DiD)[3], which was recently adopted by scholars[4], based on a different data set than us, to assess the effect of the ETS on CO_2 emissions. Specifically, using the l_1-norm row-normalized emissions data as realizations of a dependent variable Y, we performed a two-way fixed effects linear regression with multiple periods[5], whose model is defined as follows:

$$Y_{it} = \alpha + \theta_t + \eta_i + \beta_1 D_{it}^1 + \beta_2 D_{it}^2 + \varepsilon_{it}, \tag{B.1}$$

where Y_{it} is the value of row-normalized[1] (total) CO_2 emissions for the country i at time t, with $i = 1,\dots,N = 18$ and $t = 1,\dots,\tau = 17$; α is the intercept; $\theta_t \in \mathbb{R}^{\tau-1}$ and $\eta_i \in \mathbb{R}^{N-1}$ represent, respectively, the time-varying and the country-varying fixed effects (without loss of generality, respectively, with 2000 as the baseline year, and Australia as the baseline country); ε_{it} is a traditional error term, modeled as a zero mean (uncorrelated) white noise process; D^1 is a dummy variable whose realization D^1_{it} takes on the value 1 if country i is treated and time t is in phase 1 (2005, 2006, or 2007), otherwise it equals 0; and D^2 is a dummy variable whose realization D^2_{it} takes on the value 1 if country i is treated and time t is in phase 2 (2008 or later), otherwise it equals 0. The interest was on the estimates of the coefficients β_1 and β_2, which express the magnitude of the treatment in the two phases.

Results are reported in Supplementary Table B.1. According to the results, the t-student test produces a non-significant treatment effect in phase 1 and a significant treatment effect in phase 2. Because the data were normalized by row, the β_1 and β_2 coefficients needed to be re-interpreted in terms of the average percentage variation in CO_2 emissions due to a unitary increase in D^1 and D^2 (i.e., due to the treatment). Treatment in phase 1 produced a decrease of 1.6% in CO_2 emissions, while in phase 2 the estimated decrease stood at 21.9%. This latter percentage was comparable with the one we obtained in the main text using the MCFE method (19.4%).

It is worth mentioning here that the DiD method hinges on the critical assumption of parallel trend, meaning that both treated and untreated units exhibit similar trends during the pre-treatment period. However, upon visual inspection of our data, we observed that this assumption does not hold in the first and last years of pre-treatment (see Supplementary Figure B.1). Our MCFE approach provides a remedy to this issue, by not requiring this assumption. Moreover, likewise DiD, it still takes into account time-varying and country-varying fixed effects, but generalizes that method by estimating possible relationships between distinct countries and different times, modeled by the objective function of the optimization problem (1). Furthermore, recent research by[6] indicates that MCFE typically outperforms the DiD method in the presence of a ground truth.



Supplementary Figure B.1. Averages on the treated and untreated countries, respectively, of the row-normalized total CO_2 emissions before treatment: the parallel trend assumption in the pre-treatment period is violated (the slopes in the first year are quite different and in the last year are, respectively, negative/positive). The reader is referred to Supplementary Section B for an additional discussion of this figure.

Supplementary Section C Comparison with the Synthetic Control Method (SCM)

Additionally, we compared the results obtained by MCFE with the ones achieved by the Synthetic Control Method (SCM), as described by[7]. Such method was implemented using the MATLAB package "Synth", which was written specifically for that work, and is available for free at the following hyperlink: https://web.stanford.edu/~jhain/Synth_Matlab/Synth_MATLAB.zip. Without entering technical details of the method (for which the reader is referred to[7]), the basic idea of SCM, as applied to the specific data set, is to reconstruct the pre-treatment time-series of each treated country using a linear combination (with non-negative coefficients) of all the pre-treatment time-series associated with the various untreated countries, then extrapolating that linear combination to the treatment period, with the aim of predicting counterfactual values

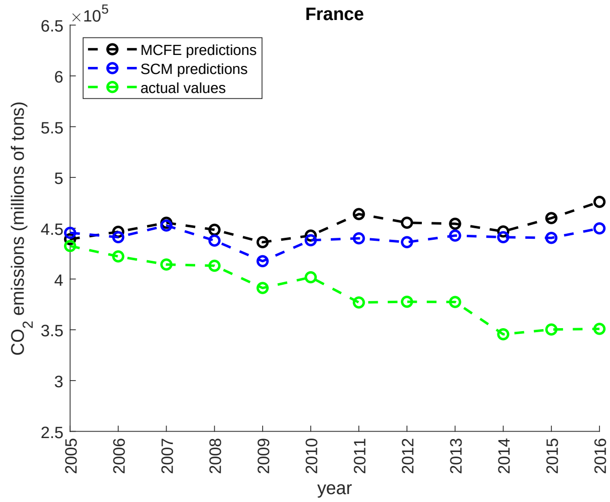
[1] For a fair comparison, the same row-normalization is considered here as for our application of the MCFE method in the main text. Since, still for fairness reasons, such row-normalization depends, for each row, on the choice of the subset of the training set associated with that row, the following results refer to the row-normalization corresponding to a specific representative random realization of the training set.

Variable	Estimated coefficient (Standard error)
(Intercept)	0.0556 (0.0014)***
factor(country) AUT	0.0007 (0.0016)
factor(country) BEL	0.0000 (0.0016)
factor(country) CAN	0.0000 (0.0013)
factor(country) DEU	0.0014 (0.0016)
factor(country) DNK	0.0092 (0.0016)***
factor(country) ESP	0.0039 (0.0016)*
factor(country) FIN	−0.0024 (0.0016)
factor(country) FRA	0.0017 (0.0016)
factor(country) GBR	0.0015 (0.0016)
factor(country) GRC	−0.0037 (0.0016)*
factor(country) IRL	0.0079 (0.0016)***
factor(country) ITA	0.0009 (0.0016)
factor(country) JPN	−0.0002 (0.0013)
factor(country) KOR	0.0017 (0.0013)
factor(country) PRT	−0.0003 (0.0016)
factor(country) SWE	−0.0019 (0.0016)
factor(country) USA	−0.0002 (0.0013)
factor(year) 2001	0.0013 (0.0013)
factor(year) 2002	0.0019 (0.0013)
factor(year) 2003	0.0029 (0.0013)*
factor(year) 2004	0.0026 (0.0013)*
factor(year) 2005	0.0031 (0.0017)
factor(year) 2006	0.0034 (0.0017)*
factor(year) 2007	0.0034 (0.0017)*
factor(year) 2008	0.0082 (0.0015)***
factor(year) 2009	0.0046 (0.0015)**
factor(year) 2010	0.0065 (0.0015)***
factor(year) 2011	0.0048 (0.0015)**
factor(year) 2012	0.0036 (0.0015)*
factor(year) 2013	0.0027 (0.0015)
factor(year) 2014	0.0010 (0.0015)
factor(year) 2015	0.0019 (0.0015)
factor(year) 2016	0.0024 (0.0015)
D^1	−0.0009 (0.0015)
D^2	−0.0121 (0.0011)***
R^2	0.7298
Adj. R^2	0.6948
Num. obs.	306

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Supplementary Table B.1. Results of the DiD two-way fixed effects linear regression with multiple periods (phase 1: 2005–2007; phase 2: 2008–2016; l_1-norm row-normalized CO_2 emissions used as realizations of the dependent variable; baseline year: 2000; baseline country: Australia). The reader is referred to Supplementary Section B for an additional discussion of this table.

for the treated country. For simplicity, here we compared the MCFE and SCM predictions by focusing on the same examples considered in Supplementary Section A. To achieve fairness in the comparison, first the SCM model was applied to the same row-normalized data considered for the MCFE application, then its predictions were de-normalized by row. Specifically, SCM was applied to estimated counterfactual values for the same treated country considered in Supplementary Figure A.1 (i.e., France), taking as input the row-normalized pre-treatment time-series associated with that country and all the row-normalized pre-treatment time-series associated with the various untreated countries (no covariates were included in the various time-series, again with the aim of achieving fairness of comparison with the MCFE application). Then, the set of SCM hyperparameters



Supplementary Figure C.1. For one MC repetition and the subsets of the test set associated with one representative treated country: predictions obtained by MCFE, predictions obtained by SCM, and corresponding actual values. The representative treated country considered matches the one of Supplementary Figure A.1.

(contained in a symmetric and positive semidefinite matrix $\mathbf{V}$, for the definition of which the reader is referred to[7]) was chosen by optimizing the mean-squared reconstruction error on the row-normalized pre-treatment time-series associated with the specific treated country. The obtained results are reported in Supplementary Figure C.1.

It is evident from Supplementary Figure C.1 that the results obtained by SCM turned out to be quite similar to the ones achieved by MCFE. However, the following are some important differences, according to which, for the specific problem under study, MCFE still looks preferable as a method for counterfactual prediction: first, the number of untreated countries (5) is so small that permutation tests such as the placebo-in-space test (see[8])[2] cannot be practically applied to SCM, to test the null hypothesis of no effect on each treated country (indeed, the resulting approximate p-value would be discretized using a too coarse grid, as explained by[9], nor refinements such as the leave-two-out in-space placebo test [3] developed in that work seem to be applicable in this case), whereas the availability of different repetitions for MCFE made it possible to perform statistical tests related to the MCFE application, as reported in the main text; SCM predicts counterfactual values for treated countries during the treatment period, by considering each such country separately and using each time only a subset of available observations, whereas MCFE predicts counterfactual values simultaneously for all the treated countries during the treatment period, using data coming from the whole portion of observed matrix of (row-normalized) CO_2 total emissions; according to[6], MCFE typically outperforms SCM in cases for which a ground truth is available.

Supplementary Section D Additional Tables

Supplementary Table D.1 reports numerical values corresponding to the years of the treatment for the plots of Figure 3.

Supplementary Table D.2 reports, for the treated countries: the sum of actual values (in thousands of tons) over the whole period of treatment covered in the database; the sum of the medians of the estimated counterfactual values (in thousands of tons) over the same period; their ratio (expressed as a percentage).

[2]Loosely speaking, this test is based on estimating effects on untreated countries as if they were treated, under the null hypothesis of no effect on each treated country.

[3]Loosely speaking, that test removes two control units at a time, instead of only one (as is usually done in the case of the traditional in-space placebo test). In this way, it is possible to increase the number of analyses used for the approximate computation of the p-value, thus reducing the granularity of its approximation. Unfortunately, for our specific case study, with only 5 available untreated countries, the resulting granularity is still too high.

Country	Quantity	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
AUT	Actual	84294	80342	79163	71406	65431	70132	68446	62079	61555	57325	62941	59850
AUT	p10	81136	80824	83153	80611	78268	81770	82394	82662	82684	81116	82873	85824
AUT	p50	81573	81261	83878	81343	81334	82843	83830	84019	84308	82696	84399	87366
AUT	p90	82235	81893	84671	82453	82820	84238	84768	85231	85447	84188	86151	90359
BEL	Actual	125516	122431	119281	118295	106258	113520	104000	101579	103016	97560	101607	100965
BEL	p10	131680	130904	134086	130434	126976	131296	131651	131170	132160	130892	133397	136306
BEL	p50	132469	131874	135153	132023	128909	134256	135350	134614	135698	134519	135728	139171
BEL	p90	134216	133170	137308	132900	131018	135848	137099	138765	138885	136409	138214	143289
DEU	Actual	980189	988849	960745	925825	853758	898560	871378	880920	896745	855312	861048	865548
DEU	p10	1045347	1040381	1068365	1037372	1006865	1043664	1052413	1052796	1054552	1041347	1061350	1091361
DEU	p50	1051333	1046341	1076412	1045108	1027503	1063086	1074788	1073229	1081853	1066958	1077962	1111233
DEU	p90	1062237	1055685	1088707	1056650	1047457	1078756	1088526	1097834	1099489	1080036	1099595	1146760
DNK	Actual	87189	104792	101229	95522	88924	87493	85286	78753	76789	71498	71812	76535
DNK	p10	81697	81286	83642	80994	78551	81528	82415	82451	82763	81484	82873	85709
DNK	p50	82261	81706	84520	81591	80613	82722	83636	83677	84445	82980	84781	87583
DNK	p90	82765	82357	85240	82646	82489	84080	84862	85327	85297	84305	86434	90403
ESP	Actual	367776	359910	374958	348386	308150	295045	296509	292142	266558	270386	287768	274796
ESP	p10	338342	337153	347069	336568	326788	341517	344051	345450	345342	338403	346103	357608
ESP	p50	340573	339232	349995	339944	340386	346317	350279	351044	351927	344953	352015	364617
ESP	p90	343371	342135	353491	344585	346534	352493	354252	356559	357113	351783	359300	377010
FIN	Actual	68615	80650	79619	62187	59478	67961	60504	54809	55753	51307	47855	51885
FIN	p10	80032	79624	82155	79329	77237	80739	81795	82159	81717	80173	82070	85151
FIN	p50	80524	80317	83149	80432	81052	81871	83050	83228	83382	81406	83727	86908
FIN	p90	81246	80990	83968	81713	82966	83465	83954	84972	85485	83369	85743	90394
FRA	Actual	432515	422251	414279	413064	391160	401726	376896	377612	377373	345587	350293	350911
FRA	p10	444826	442692	453999	441261	428618	444230	446344	445383	447581	442716	451217	462512
FRA	p50	447567	445483	457148	445635	436805	453469	457631	455860	459269	454497	459075	471591
FRA	p90	453074	449380	463751	449369	444135	459232	463288	468319	468700	460606	466904	485965
GBR	Actual	633973	623064	615083	600149	545276	563004	524184	537303	524550	488332	474426	450554
GBR	p10	632730	630215	645846	628763	614010	633610	636201	634010	637293	629346	645014	660660
GBR	p50	636493	634069	649782	635880	624395	648396	652967	649607	653636	647008	654569	672970
GBR	p90	644749	640484	660415	640177	634416	655784	660195	667576	667088	654928	665946	692274
GRC	Actual	139792	138028	142127	111113	104341	97343	94532	91418	81723	78656	74962	71373
GRC	p10	137630	137027	140689	136774	132894	137889	138741	138807	139019	137206	140035	144219
GRC	p50	138389	137829	141727	137893	135965	140356	141686	141459	142378	140423	142231	146767
GRC	p90	139766	138968	143369	139233	138449	142249	143336	144683	144644	142193	144910	151309
IRL	Actual	47425	46833	46807	46981	41811	48998	45417	46178	43933	44079	46313	49104
IRL	p10	45297	45114	46161	45107	44286	45556	45650	45521	45643	45038	46302	47411
IRL	p50	45587	45471	46484	45696	44946	46704	47005	46755	46904	46511	46902	48217
IRL	p90	46274	45988	47410	46015	45613	47371	47562	48103	48021	47116	47745	49595
ITA	Actual	513883	509526	498800	487925	428896	437579	425745	402053	372699	356597	365070	359005
ITA	p10	504903	502894	516444	502101	488423	506476	509105	509706	510230	503123	514485	530021
ITA	p50	507711	505847	519772	506611	500059	515866	520460	519260	522391	515264	522430	539149
ITA	p90	512811	509925	526140	511112	508878	522339	525984	531113	530730	522289	531737	555476
PRT	Actual	71982	67336	64419	62595	58475	53819	52720	51190	49283	49364	53440	51532
PRT	p10	69364	68969	70951	68795	66834	69488	70081	70317	70495	69129	70481	72609
PRT	p50	69721	69430	71368	69417	68674	70808	71663	71734	72071	70835	71531	73809
PRT	p90	70554	70067	72322	70288	69946	72033	72443	73169	73120	71935	73051	76291
SWE	Actual	65943	63968	62716	60166	56164	61592	57081	54286	52187	51176	51454	51765
SWE	p10	73071	72607	74581	72238	70189	72491	73357	73437	73816	72669	73856	75842
SWE	p50	73501	73056	75191	72711	71358	73846	74915	74863	75545	74484	75146	77316
SWE	p90	74191	73570	75968	73741	72870	75186	75967	76656	76805	75583	76787	79952

Supplementary Table D.1. Comparison between actual values (in thousands of tons) and estimated counterfactual values (in thousands of tons) for the treated countries in the years of treatment. In the table, p10, p50 and p90 stay, respectively, for the 10^{th}, the 50^{th} and the 90^{th} percentile.

	Sum of Actual (period: 2005–2016)	Sum of Estimated (period: 2005–2016)	Sum of Actual/Sum of Estimated (expressed as a percentage)
Austria	822964	998851	82.39
Belgium	1314029	1609764	81.63
Germany	10838877	12795806	84.71
Denmark	1025822	1000515	102.53
Spain	3742384	4171282	89.72
Finland	740623	989045	74.88
France	4653667	5444030	85.48
Great Britain	6579898	7759771	84.80
Greece	1225408	1687103	72.63
Ireland	553879	557183	99.41
Italy	5157778	6194820	83.26
Portugal	686155	851060	80.62
Sweden	688498	891933	77.19

Supplementary Table D.2. For the treated countries: sum of actual values (in thousands of tons) over the whole period of treatment covered in the database (Sum of Actual), sum of the medians of the estimated counterfactual values (in thousands of tons, median is used) over the same period (Sum of Estimated), and the ratio Sum of Actual/Sum of Estimated, expressed as a percentage.

Similarly, Supplementary Table D.3 reports, for the treated countries: the sum of actual values (in thousands of tons) over the first period of treatment; the sum of the medians of the estimated counterfactual values (in thousands of tons) over the same period; their ratio (expressed as a percentage).

	Sum of Actual (period: 2005–2007)	Sum of Estimated (period: 2005–2007)	Sum of Actual/Sum of Estimated (expressed as a percentage)
Austria	243799	246712	98.82
Belgium	367228	399496	91.92
Germany	2929783	3174086	92.30
Denmark	293210	248487	118.00
Spain	1102644	1029799	107.07
Finland	228884	243990	93.81
France	1269045	1350198	93.99
Great Britain	1872120	1920344	97.49
Greece	419947	417945	100.48
Ireland	141065	137542	102.56
Italy	1522209	1533330	99.27
Portugal	203737	210518	96.78
Sweden	192627	221748	86.87

Supplementary Table D.3. For the treated countries: sum of actual values (in thousands of tons) over the first period of treatment (Sum of Actual), sum of estimated counterfactual values (in thousands of tons, median is used) over the same period (Sum of Estimated), and the ratio Sum of Actual/Sum of Estimated, expressed as a percentage.

Supplementary Table D.4 expresses the reduced damage in monetary terms per capita (in USD 2020 per capita) for the treated EU countries during the whole treatment period covered in the database.

	Reduced damage in monetary terms per capita (in USD 2020 per capita)											
Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Austria	-61	21	105	221	353	281	339	481	496	549	459	583
Belgium	123	166	276	237	388	352	525	550	542	610	560	624
Germany	160	129	260	269	392	372	469	442	425	483	491	552
Denmark	-168	-785	-566	-469	-278	-159	-55	163	252	376	422	357
Spain	-115	-86	-102	-34	129	204	213	233	339	297	256	357
Finland	420	-12	123	635	748	480	774	971	940	1020	1211	1179
France	44	68	124	94	130	147	229	220	230	304	302	335
Great Britain	8	33	105	107	235	252	377	326	373	454	512	627
Greece	-24	-3	-7	447	527	716	786	838	1023	1049	1150	1294
Ireland	-82	-59	-14	-53	128	-93	64	23	119	97	23	-35
Italy	-20	-12	66	59	223	244	295	364	460	483	479	550
Portugal	-40	37	122	120	179	297	332	361	403	382	323	399
Sweden	155	185	252	252	302	242	349	400	450	445	447	476

Supplementary Table D.4. Reduced damage in monetary terms per capita (expressed in USD 2020 per capita) for the treated EU countries in the treatment period covered in the database.

References

1. Metulini, R., Biancalani, F., Gnecco, G. & Riccaboni, M. Assessing the performance of nuclear norm-based matrix completion methods on CO_2 emissions data. In *Book of the Short Papers - SIS 2023 - Statistical Learning, Sustainability and Impact Evaluation (Eds. F. M. Chelli, M. Ciommi, S. Ingrassia, F. Mariani, M. C. Recchioni)*, 776–781 (Pearson, 2023).
2. Capponi, A. & Stojnic, M. Exact error in matrix completion: Approximately low-rank structures and missing blocks (2023). Https://arxiv.org/abs/2401.00578v1.
3. Rubin, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66**, 688 (1974).
4. Bordignon, M. & Gamannossi degl'Innocenti, D. Third time's a charm? Assessing the impact of the tihird phase of the EU ETS on CO_2 emissions and performance. *Sustainability* **15**, 6394 (2023).
5. Callaway, B. & Sant'Anna, P. H. C. Difference-in-differences with multiple time periods. *J. Econom.* **225**, 200–230 (2021).
6. Athey, S., Bayati, M., Doudchenko, N., Imbens, G. & Khosravi, K. Matrix completion methods for causal panel data models. *J. Am. Stat. Assoc.* **116**, 1716–1730 (2021).
7. Abadie, A., Diamond, A. & Hainmueller, J. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *J. Am. Stat. Assoc.* **105**, 493–505 (2010).
8. Yan, G. & Chen, Q. Synth2: Synthetic control method with placebo tests, robustness test, and visualization. *The Stata J.* **23**, 597–624 (2023).
9. Lei, L. & Sudijono, T. Inference for synthetic controls via refined placebo tests (2024). Https://arxiv.org/abs/2401.07152.