Zero fare, cleaner air? The causal effect of Luxembourg’s free public transportation policy on carbon emissions

Tobias Eibinger[1] & Sachintha Fernando[2]

[1]University of Graz, AT

[2]University of Halle-Wittenberg, DE

EEA Rotterdam

August 27, 2024

Introduction

- In March 2020, Luxembourg became the first country to implement ***free public transport for everybody across all modes***

  ⟶ Intention: reduce car density and associated negative externalities

- **RQ:** What are the effects of this policy on transport CO2 emissions?

- Identification challenges:
  1. Luxembourg is special
  2. COVID-19 - and associated variation in mobility patterns

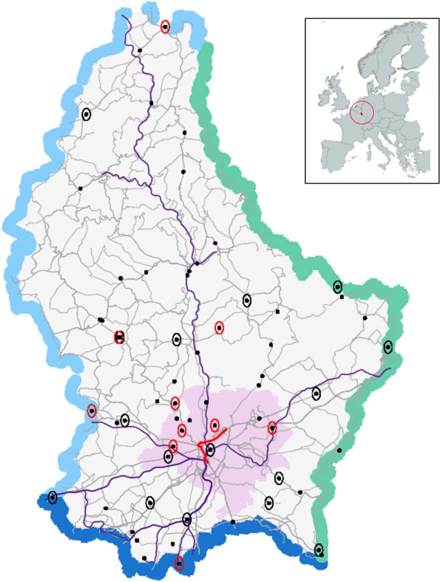
- Data on CO2 emissions: IPCC-sector 1.A.3.b Grided CO2 emission data from EDGAR

- Unit of analysis: NUTS 2 level

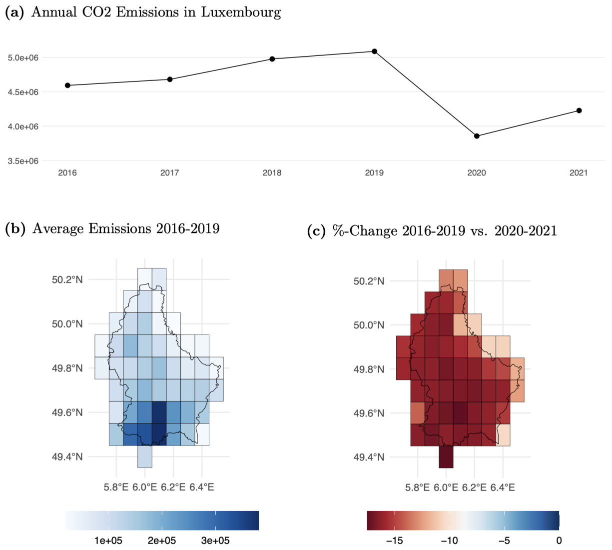
- Method: Synthetic Difference in Differences (SDID) (Arkhangelsky et al., 2021)

- **Results**: ATT of around -6.1% CO2 road transport emissions

Public Transport in Luxembourg



Evolution of road transport emissions from LU



Luxembourg is special

- Size: 2,586.4 km^2
- Population: $\approx$ 660,000
- GDP/capita: $\approx$ 140,000 USD (highest in EU)
- Car density: $\approx$ 700 cars per 1,000 inhabitants (highest in EU)

Challenge:

- difficult to meet the parallel trend assumption for DiD
- SC assumes comparable pre-treatment levels

Solution:

- Unit of analysis: NUTS 2 level
- Use SDID, which combines characteristics of both SC and DID. Advantage: Does not assume comparable levels in any stage

Potential Confounding

Main threat: COVID-19

Has mobility behavior in LU changed differently compared to other regions?

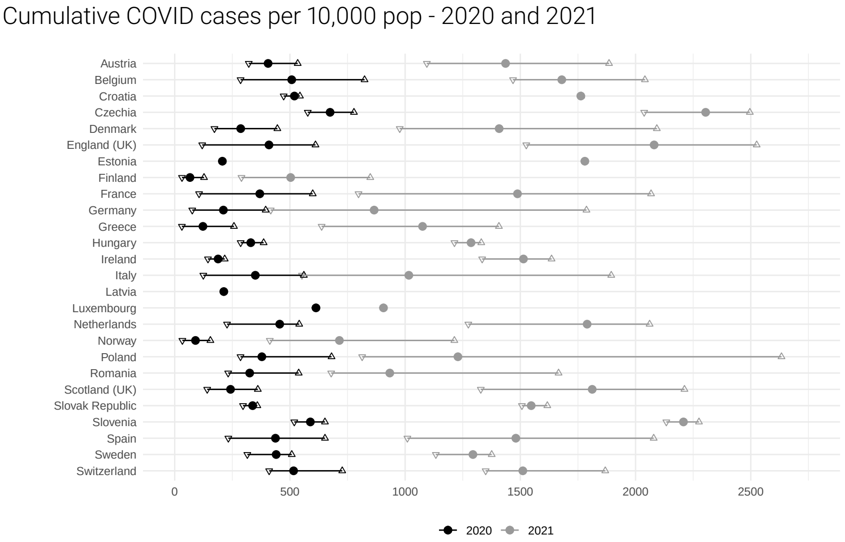
- COVID-19 cases
- Commuting
- Working from home

Other threats

- Fuel prices, energy efficiency of new vehicles, freight volume

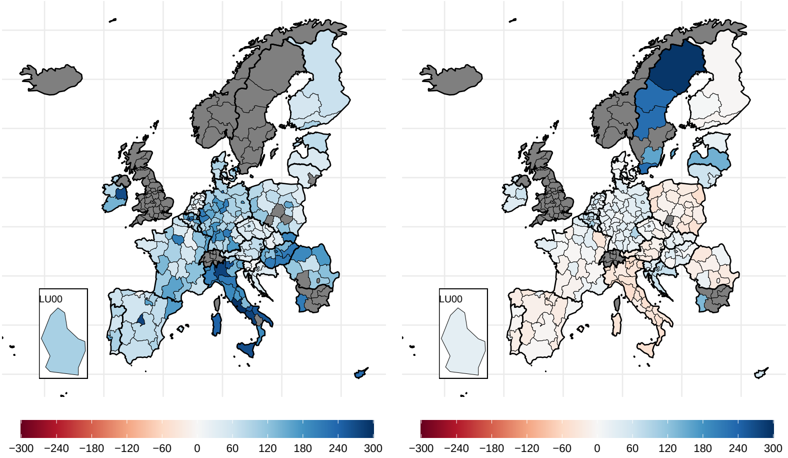
Consider bad comparisons and spillovers

COVID - 19


Cumulative COVID cases per 10,000 pop - 2020 and 2021
Austria
Belgium
Croatia
Czechia
Denmark
England (UK)
Estonia
Finland
France
Germany
Greece
Hungary
Ireland
Italy
Latvia
Luxembourg
Netherlands
Norway
Poland
Romania
Scotland (UK)
Slovak Republic
Slovenia
Spain
Sweden
Switzerland
0
500
1000
1500
2000
2500
2020
2021

COVID - 19 cont ...

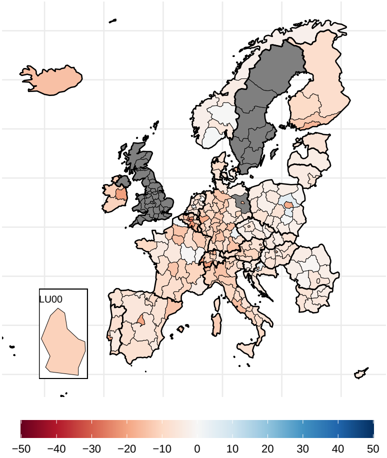
Change (%) people usually WFH with workplace in NUTS2 region and residency in the same country: 2019-2020 and 2020-2021

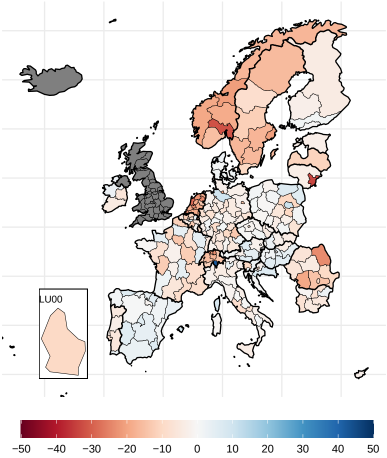


COVID - 19 cont...

Change (%) overall commuting inflow irrespective of residency:

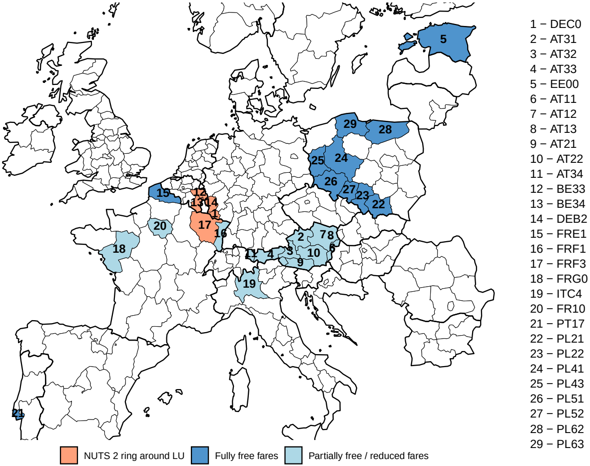
2019-2020 and 2020-2021





Bad comparisons and spillovers (2016-2021)

- Exclude NUTS 2 ring around LU
- Exclude regions that introduced fully free fares during sample period (2016-2021)



Synthetic DiD

We apply syhthetic DiD to **emissions adjusted** for covariates

Synthetic DiD combines features of both DiD and SC methods.

- Like DiD, it is invariant to additive unit-level shifts
- Like SC, it weighs and matches pre-treatment trends to reduce reliance on parallel-trends assumption

Synthetic DiD re-weighs **both units and time periods.**

- Unit weights to match pre-treatment trends between exposed and unexposed units.
- Time weights assign higher weights to pre-treatment time periods that are more similar to post-treatment time periods for unexposed units.

Synthetic DiD

$$\left(\widehat{\tau}^{sdid}, \widehat{\mu}, \widehat{\alpha}, \widehat{\beta}\right) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \widehat{\omega}_i^{sdid} \widehat{\lambda}_t^{sdid} \right\}$$

▸ weights

Synthetic Control

$$\left(\widehat{\tau}^{sc}, \widehat{\mu}, \widehat{\beta}\right) = \underset{\tau,\mu,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \beta_t - W_{it}\tau)^2 \widehat{\omega}_i^{sc} \right\}$$

DiD

$$\left(\widehat{\tau}^{did}, \widehat{\mu}, \widehat{\alpha}, \widehat{\beta}\right) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \right\}$$

Placebo Inference

$\widehat{\tau}^{sdid}$ is asymptotically normal $\longrightarrow$ conventional CIs can be used if the asymptotic variance can be consistently estimated. $\tau \in \widehat{\tau}^{sdid} \pm z_{\alpha/2}\sqrt{\widehat{V}_\tau}$.

With $N_{tr} = 1$, we can use **placebo based inference**:

- Replace the exposed unit with unexposed units
- Randomly assign those units to a placebo treatment
- Compute a placebo ATT
- Repeat many times to obtain a vector of placebo ATTs

Event-study inference can be conducted by estimating:

$$d_t = (\bar{Y}_t^1 - \bar{Y}_t^0) - (\bar{Y}_{base}^1 - \bar{Y}_{base}^0).$$

Confidence bands around these estimates can be generated with a placebo-based approach (Arkhangelsky et al., 2021; Clarke et al., 2023). es-placebo

Handling covariates

Handling covariates in this setting is treated as a **pre-modelling approach**. Model with fixed effects is estimated only for **control regions** as suggested by Kranz (2022):

$$Y_{it}^{co} = \alpha_i + \gamma_t + X_{it}^{co}\beta + u_{it}, \quad (1)$$

$$\widehat{Y}_{it}^{adj} = Y_{it} - X_{it}\hat{\beta}. \quad (2)$$

- log of real GDP/CAP (regional)
- asinh of daily COVID cases (regional)
- asinh commuting inflow (regional) scatter
- asinh working from home (regional)
- log of emission intensity of new vehicles (national)
- diesel and super prices in real terms (national)
- log of total freight goods loaded (regional)

Projected

Three specifications

1. No covariates.
2. Adjusted for COVID-19 related covariates.

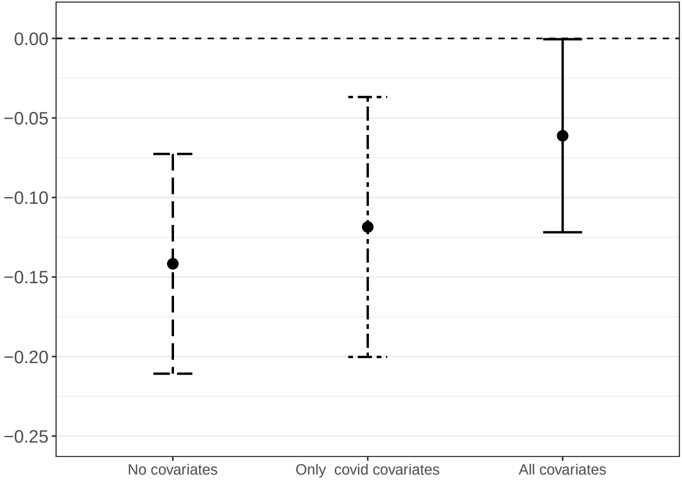
$$
\begin{aligned}
log(CO2/cap)_{it}^{co} =& \alpha_i + \gamma_t + \beta_1 asinh(cases)_{it}^{co} + \beta_2 asinh(nvrwfh)_{it}^{co} + \\
& \beta_3 asinh(wfh)_{it}^{co} + u_{it},
\end{aligned}
$$

3. Full set of covariates (**main specification**)

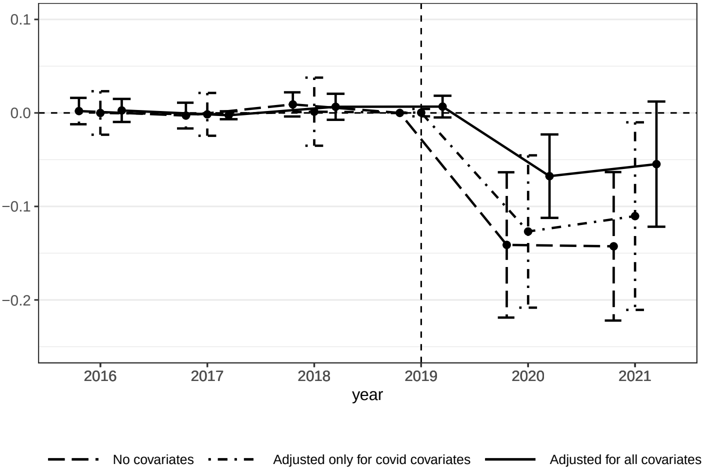
$$
\begin{aligned}
log(CO2/cap)_{it}^{co} =& \alpha_i + \gamma_t + \beta_1 asinh(cases)_{it}^{co} + \beta_2 asinh(nvrwfh)_{it}^{co} + \\
& \beta_3 asinh(wfh)_{it}^{co} + \beta_4 log(gdp)_{it}^{co} + \beta_5 log(ei)_{it}^{co} + \\
& \beta_6 diesel_{it}^{co} + \beta_7 petrol_{it}^{co} + \beta_8 log(frt)_{it}^{co} + u_{it}.
\end{aligned}
$$

Results - ATT

- ATT of -6.1% (specification with all covariates) ▸ Unit weights ▸ Time weights



Results - Event Study



Robustness

- Pre-trend with sdid weights ▸ pre-trend
- In-time placebo: 2018 treatment ▸ in-time placebo
- Different specificaions ▸ other specs
  - → exclude freight
  - → exclude wfh
  - → exclude commuting
- Restricted sample: Drop all regions that introduced free / reduced fares in our sample period ▸ limited sample
- Relative fuel prices (fuel tourism) ▸ rel fuel fig ▸ rel fuel tab

Back of the envelope calculations

1. Effect size discussion

- Following Bigi et al. (2023), let us assume a modal split for private vehicles and public transport of around 80% and 15%
- Assume emission reduction is due to a modal change from private vehicles to public transport
  $\longrightarrow$ Estimated increase of public transport: $(\hat{\tau}80\%/15\%) \approx 32\%$.
- In line with LuxMobile survey: 30% increase in public transport usage due to the free-fare policy

2. Marginal abatement cost of carbon

- Foregone revenue from ticket sales of around 41 Mio. Euros
- Compare to CO2 emissions abated according to our estimates: $(CO2_t^{pre}\hat{\lambda}_t)\hat{\tau}$
  $\longrightarrow$ EUR 140 per tonne of carbon.

Recap

- We study the *effect of free public transport on transport CO2 emissions*.
- We use SDID to create a comparable counterfactual to Luxembourg
- We control for potential confounders
  - → COVID, working from home, commuting, fuel prices, emission intensity of new vehicles, freight transport
- We estimate an *ATT of around* -6.1%
- Results hold against robustness checks
- Framework to study other policies during Covid

Thank you

Questions, comments, and suggestions are welcome

tobias.eibinger@uni-graz.at

sachintha.fernando@wiwi.uni-halle.de

References

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159

Bigi, F., Schwemmle, N., & Viti, F. (2023). Evaluating the impact of free public transport using agent-based modeling: The case-study of luxembourg.

Clarke, D., Pailañir, D., Athey, S., & Imbens, G. (2023). Synthetic difference in differences estimation. https://arxiv.org/abs/2301.11859

Kranz, S. (2022). *Synthetic difference-in-differences with time- varying covariates. technical report*. https://github.com/skranz/xsynthdid/blob/main/paper/synthdid%20with%20covariates.pdf

Unit weights are computed to align pre-treatments trends between treated and control units:

$$\left(\widehat{\omega}_0, \widehat{\omega}^{sdid}\right) = \underset{\omega_0 \in \mathbb{R}, \omega \in \Omega}{\arg\min} \sum_{t=1}^{T_{pre}} \left(\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2 + \zeta^2 T_{pre} ||\omega||_2^2.$$

Time weights are computed to align pre- and post-treatment periods of control units:

$$\left(\widehat{\lambda}_0, \widehat{\lambda}^{sdid}\right) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\arg\min} \sum_{i=1}^{N_{co}} \left(\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} \right)^2 + \zeta^2 N_{co} ||\lambda||^2.$$

- N_{co} and N_{tr} are the number of untreated and treated units.
- T_{pre} and T_{post} are the number of pre- and post-treatment periods.
- ζ is a regularization parameter to increase dispersion and ensure unique weights as defined in Arkhangelsky et al. (2021)

back

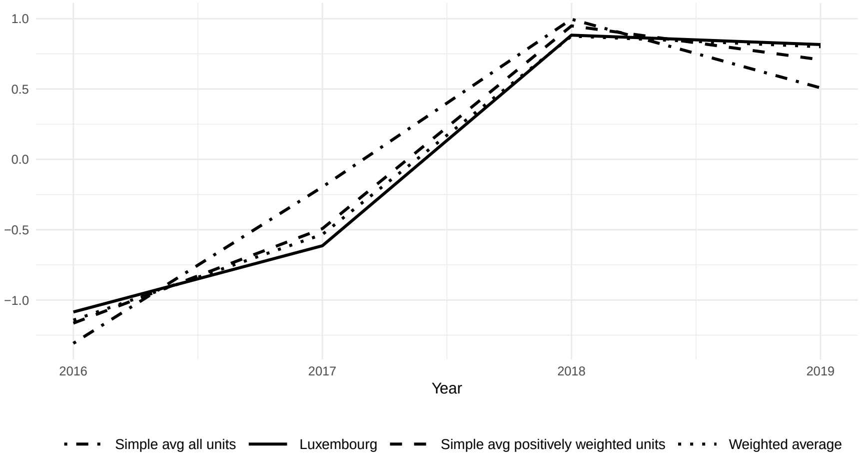
Covariates - Projected

	(1)		(2)	
	Estimate	Std. Error	Estimate	Std. Error
asinh(cases)	-0.0284***	(0.0049)	-0.0119	(0.0072)
asinh(nvrwfh)	0.0789***	(0.0264)	0.1217**	(0.0480)
asinh(wfh)	-0.0148**	(0.0062)	-0.0459***	(0.0101)
log(gdp)	0.3613***	(0.0731)		
log(ei)	0.2219***	(0.0418)		
diesel	-0.7463***	(0.0919)		
petrol	0.2765**	(0.113)		
log(frt)	0.0148	(0.0097)		
Obs	816		816	
N	136		136	
T	6		6	

Notes: Dependent variable is $lco2cap$, standard errors are clustered at the regional level.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

back

Trend comparisons - normalized



back

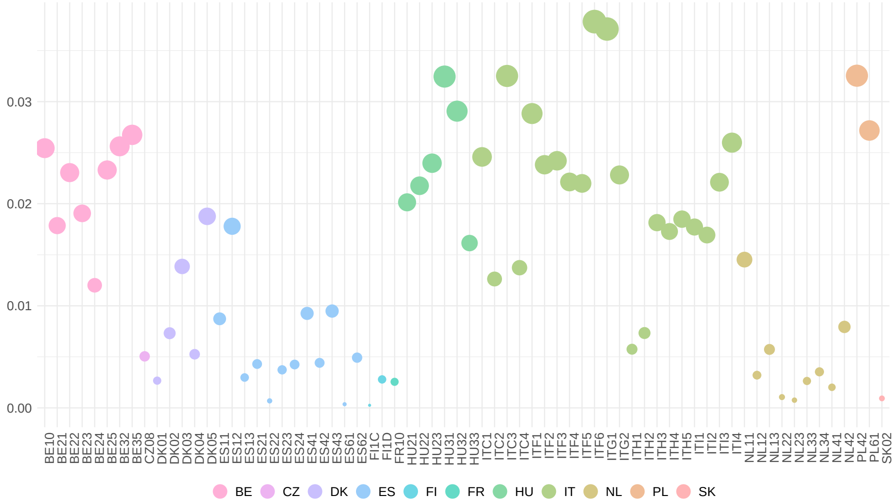
Placebo Inference

Confidence bands around the estimated d_t's are generated with a placebo-based approach in the following sequence:

1. Exclude the treated unit (in our case Luxembourg) from the sample
2. Randomly assign treatment to a unit (from the remaining units, which are all controls units)
3. Calculate the outcome adjusted for covariates, i.e., $\hat{Y}_{it}^{adj}$.
4. Compute d_t and store the result
5. Repeat 2-4 many times (e.g., 1,000 times)
6. Obtain the 5% quantile from the sample distribution of the stored results for each time period t.

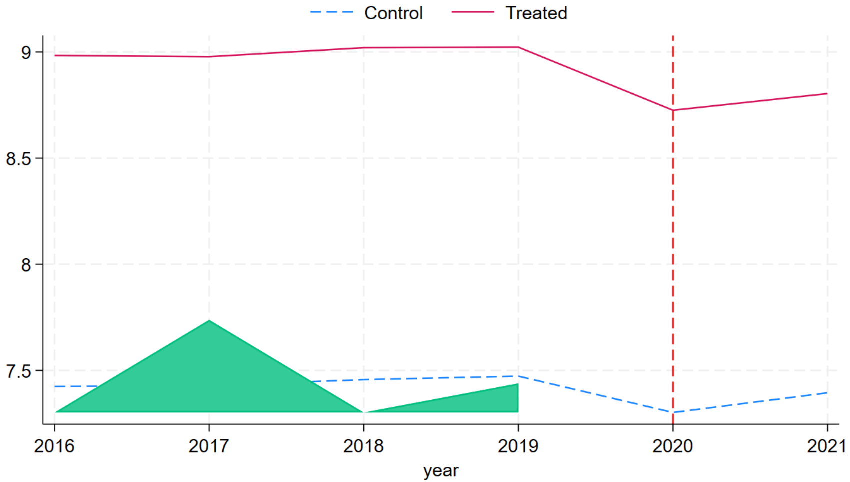
back

Unit weights - all covariates



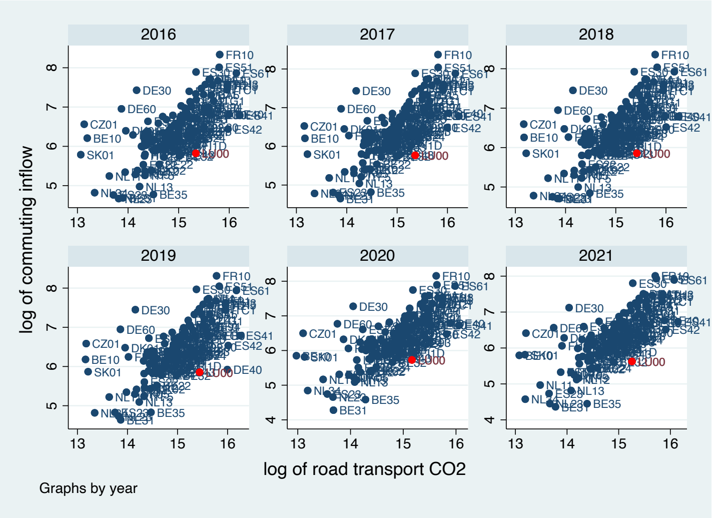
back

Time weights - all covariates



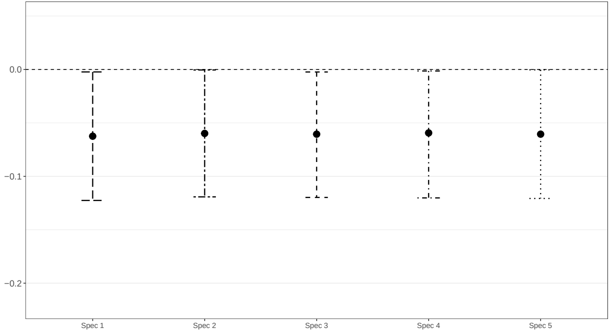
back

Commuting inflow scatter plot



▸ back

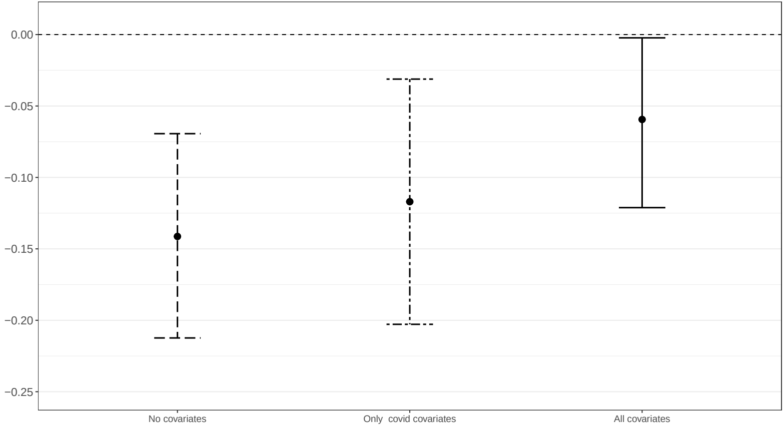
Robustness tests - ATTs across different specifications



Notes: *Spec 1* excludes controls for freight transport; *Spec 2* excludes controls for working from home; *Spec 3* excludes controls for both freight and working from home, *Spec 4* excludes controls for commuting (never working from home); *Spec 5* excludes controls for both freight and commuting

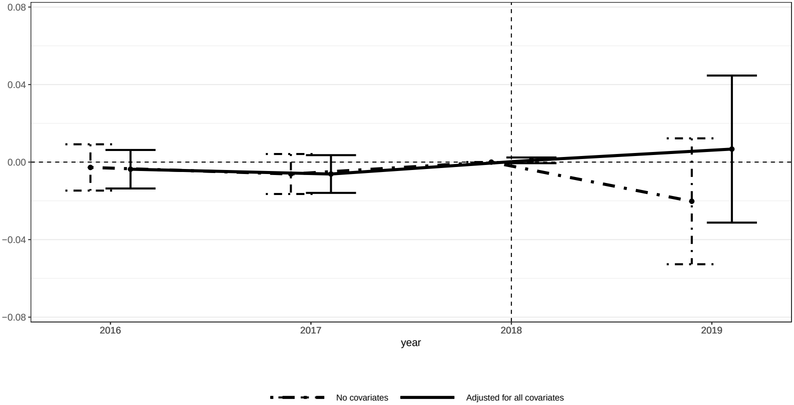
▸ back

ATTs using restricted sample



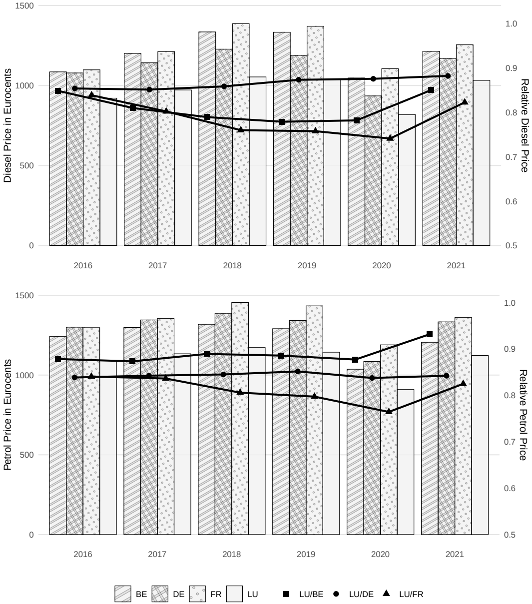
back

In-time placebo



back

Relative fuel prices



back

Relative fuel prices

Table: Pre- and post-treatment averages of relative fuel prices for Luxembourg

	Diesel		Petrol	
	Pre-Avg	Post-Avg	Pre-Avg	Post-Avg
BE	0.8010	0.8186	0.8765	0.9065
DE	0.8575	0.8794	0.8448	0.8401
FR	0.7892	0.7844	0.8253	0.7965

Note: Relative fuel prices of LU with respect to its neighboring countries. Pre-Avg are relative fuel prices based on time-weighted pre-treatment fuel prices, where time weights are taken from the SDiD main specification. Post-Avg are relative fuel prices based on post-treatment fuel prices.

back