nature energy

ARTICLES

https://doi.org/10.1038/s41560-022-01095-6

Attributing agnostically detected large reductions in road CO_2 emissions to policy mixes

Nicolas Koch[1,2,3] ✉, Lennard Naumann[1,4], Felix Pretis[5,6], Nolan Ritter[1,2] and Moritz Schwarz[6,7]

Policymakers combine many different policy tools to achieve emission reductions. However, there remains substantial uncertainty around which mixes of policies are effective. This uncertainty stems from the predominant focus of ex post policy evaluation on isolating effects of single, known policies. Here we introduce an approach to identify effective policy interventions in the EU road transport sector by detecting treatment effects as structural breaks in CO_2 emissions that can potentially occur in any country at any point in time from any number of a priori unknown policies. This search for 'causes of effects' within a statistical framework allows us to draw systematic inference on the effectiveness of policy mixes. We detect ten successful policy interventions that reduced emissions between 8% and 26%. The most successful policy mixes combine carbon or fuel taxes with green vehicle incentives and highlight that emissions reductions on a magnitude that matches the EU zero emission targets are possible.

Although ever more countries commit to net-zero greenhouse gas emissions, it remains unclear how to achieve them. Canada, the European Union, Japan, New Zealand and the United Kingdom have passed net-zero emissions targets into law while others such as China and the United States made similar pledges in official policy documents[1]. As a major emitter, the transportation sector is indispensable for achieving net-zero emissions. Globally, it emitted nearly 8.5 Gt CO_2 in 2019 or one-quarter of all GHG emissions, while the International Energy Agency suggests that annual global sector emissions need to fall below 1 Gt CO_2 by 2050 to reach net-zero overall emissions[1]. So far, the sector has proven resilient to emissions reduction efforts, especially in road transport[2].

The question how to best achieve net-zero emission targets triggers fierce policy debates about the appropriate means as exemplified in the context of the European Green Deal. To become the first climate-neutral continent by 2050, the European Union is currently revising its climate, energy and transport-related legislation under the so-called 'Fit-for-55 package'. One building block is the ambitious increase of EU member states' emissions reduction targets for 2030 from 30% to 40% compared with 2005 under the EU Effort Sharing Regulation (ESR). Under the ESR, each EU member state must meet binding annual emissions reduction targets for the agriculture, buildings and transport sectors by implementing national policies. Yet, according to the latest national projections available, most EU members will miss their targets pursuing current policy instruments. For transport specifically, emissions under the existing policies are projected to be at nearly the same level in 2030 as they were in 2020[3,4]. Thus, the ESR exerts considerable pressure on EU Member States to strengthen climate policies in transport. Policymakers have to choose from myriads of promising policies to achieve emissions reductions. There is a controversial policy discussion about whether the ambitious climate targets are best achieved by using a policy mix that emphasizes tax policies, such as carbon taxes, or green spending, such as electric vehicle subsidies, or command-and-control measures, such as speed limits and efficiency labels[5,6].

There remains substantial empirical uncertainty around which policy mixes are effective in actually achieving the objective they were designed for. Part of this uncertainty remains because empirical policy evaluation in the existing literature predominantly focuses on evaluating single, known interventions in isolation by posing the forward causal question of what happens as a consequence of a particular policy[7–10]. This 'effects of causes' approach[11] runs the risk of missing a priori unknown or underappreciated interventions. It also requires a context that allows for isolating a single policy's effects from simultaneously implemented and potentially confounding ones. Such contexts are rare because policymakers routinely legislate mixes of many interventions simultaneously[2,12]. When having to choose from many interacting available policy interventions, it can, however, be more intuitive to ask a reverse causal question looking for 'causes of effects'[11] to find what caused reductions in emissions (rather than whether a single policy is effective). Such a question is highly relevant to identify either unknown but effective policies or, more importantly, effective mixes of interacting policy interventions. However, in terms of technical implementation, it is less obvious how this kind of question may be tackled.

Here we introduce an approach to implement and answer the reverse causal question of 'What reduced CO_2 emissions?' in the EU road transport sector between 1995 and 2018 by first detecting substantial changes in emissions relative to a control group using machine learning and subsequently attributing them to likely causes such as single or interacting policy interventions. Because detection is separate from policy attribution, our approach neither requires any a priori knowledge of reductions in emissions, nor does it require a priori knowledge of the number of policies that caused these. Therefore, we are able to identify previously unknown policies or policy *mixes* that effectively reduce CO_2 emissions. While the EU transport sector is a policy-relevant test bed, our approach is readily applicable in many other contexts.

[1]Mercator Research Institute on Global Commons and Climate Change (MCC), Berlin, Germany. [2]Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany. [3]IZA Institute of Labor Economics, Bonn, Germany. [4]WZB Berlin Social Science Center, Berlin, Germany. [5]Department of Economics, University of Victoria, Victoria, British Columbia, Canada. [6]Nuffield College, University of Oxford, Oxford, United Kingdom. [7]Smith School of Enterprise and the Environment, University of Oxford, Oxford, United Kingdom. ✉e-mail: koch@mcc-berlin.net

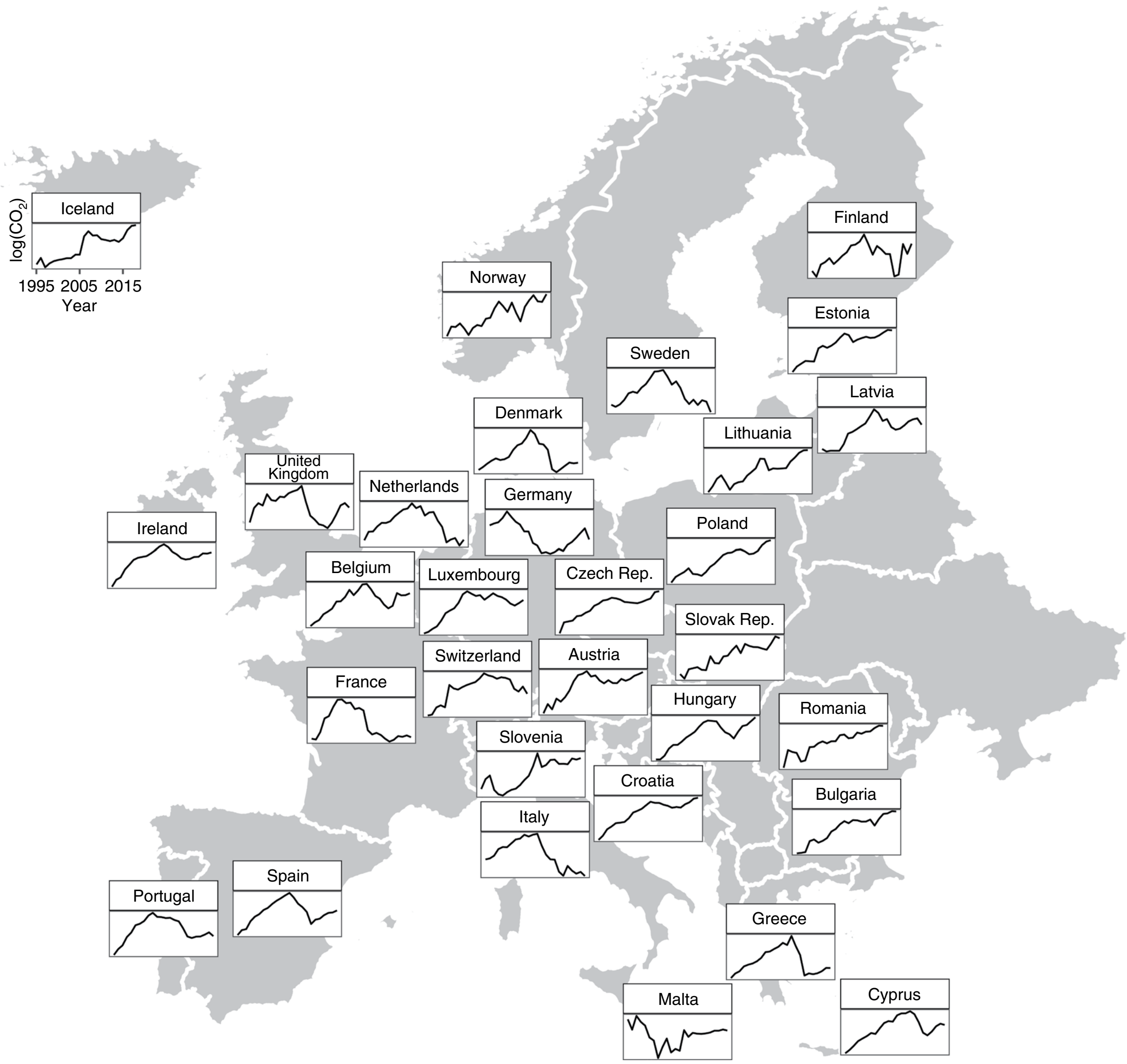

Fig. 1 | Emissions in road transport in Europe. The natural logarithm of CO_2 emissions (log(CO_2); relative) between 1995 and 2018 by country. Please refer to Iceland for year indicators on the horizontal axis. The *y* axis indicates log(CO_2). Rep., Republic. Background map from http://www.efrainmaps.es.

We produce three key findings. First, we detect ten successful policy interventions that reduced emissions between 8% and 26%. We link all these reductions to at least one tax that increases driving costs; we link seven breaks to carbon taxes, four breaks to fuel taxes and three to road tolls. Second, we link eight of the ten breaks to policy mixes that combine the aforementioned taxes with either CO_2-based vehicle taxes or subsidies for low-emissions vehicles. Third, we link the breaks with the highest level of confidence and the greatest effect sizes of up to 26% to increases of existing but moderate carbon or fuel taxes. Altogether, the ten policy interventions we identified between 1995 and 2018 reduced emissions in the EU-15 (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom) by up to 35.9 Mt CO_2. In comparison, the current ESR requires a reduction of 480 Mt CO_2 for the same region between 2021 and 2030. Even if we conservatively assume that agriculture, buildings and transport contribute in equal measure to these reductions, the transport policies implemented to date seem rather inadequate—even more so when we account for the imminent tightening of the ESR targets proposed in the EU's Fit-For-55 package. At the same time, the relative reductions of up to 26% for certain breaks indicate considerable potential for future reductions. The most successful intervention implemented in Finland in 2000 (−17%), Sweden in 2001 (−11%), Ireland in 2011 (−13%) and Luxembourg in 2015 (−26%) combine increasing carbon or fuel taxes to curb mileage with complementary financial incentives to support the transition to greener vehicles.

Using break detection to assess policies

Existing ex post policy evaluations predominantly focus on the forward causal question of 'What happens as a consequence of a particular known policy?' It is reasonably straightforward to evaluate these with time-tested, quasi-experimental tools from programme evaluation, ranging from difference-in-differences[13,14] and matching[15] to synthetic control methods[9,10,16]. However, drawing systematic inference is difficult because the available evidence is scattered across countries and policies and because the study of 'effects of causes' runs the risk of missing effective but unknown interventions or those falsely deemed ineffective. Moreover, the forward causal approach needs to ensure that treatments are independent and unconfounded, which is challenging because policymakers routinely implement mixes of many simultaneous interventions with common goals.

It is less obvious how to answer reverse causal questions such as 'What has reduced emissions?'. As Gelman and Imbens[11] put it: 'Reverse causal reasoning is different; it involves asking questions

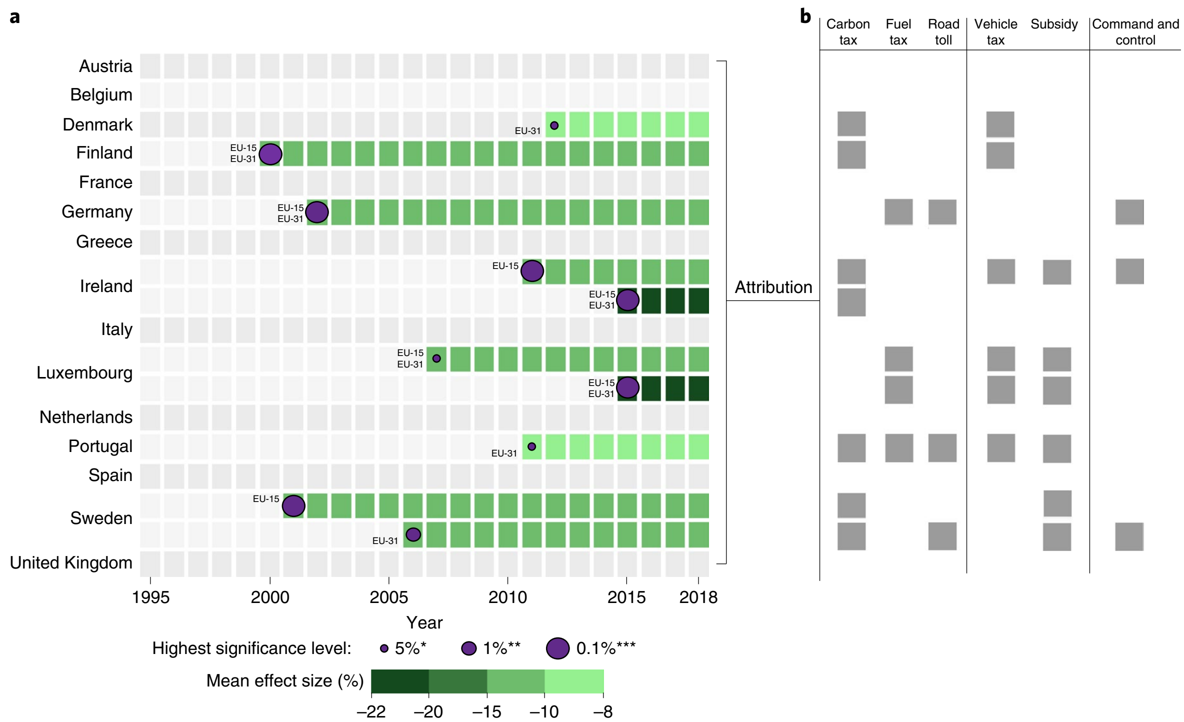

Fig. 2 | Detected breaks in road CO_2 emissions and their attribution. a, Significant breaks in CO_2 emissions. The markers' sizes indicate the highest significance level at which a break was detected. ***, **, and * indicates statistical significance at the 0.1%, 1%, and 5% level, respectively. The colour indicates effect sizes in percent. We find, at most, two detected breaks in any given country (this is an estimation result and not imposed by our Methods, which do not impose an upper limit on the number of detected breaks). Those countries occupy two lines while all others occupy a single line. **b**, The types of policy that caused the reductions.

and searching for new variables that might not yet even be in our model.' We formalize this approach by expanding on the idea of '*searching for new variables*' and placing reverse causal analysis into the domain of variable selection, and more specifically break detection. We tackle the question of 'What reduced emissions?' by identifying notable reductions in CO_2 emissions, which we identify as structural breaks. We use familiar two-way fixed effects (TWFE) panel estimators to detect these breaks and estimate a separate treatment effect for each identified break in each country (Methods and refs. [17,18] for the relevant R package 'getspanel'). Our approach identifies country- and time-specific treatment effects on the treated by detecting breaks for each policy and treated country. It reveals when an emissions break occurs within an approximate margin of error. Any policy implemented within this margin potentially caused the break. We place no restriction on the number of potential treatments nor do we impose a minimum break (that is, treatment) length. A causal interpretation rests on the assumption that there were no other influencing factors than the attributed policies themselves.

The idea of scrutinizing data for structural breaks is firmly established in the time series literature on policy evaluation (for example, ref. [19] on the Montreal Protocol, ref. [20] on UK CO_2 emissions or ref. [21] on homicides). However, time series methods lack control groups, making causal interpretations difficult. The combination of conservative significance levels (to control the false positive rate of detection) and the use of control groups in the panel setting give the reverse causal approach credibility and reduces the risk of spuriously identifying false positive results. We propose the approach to complement the traditional forward causal analysis. While the latter excels at recovering causal effects of known individual policies, the proposed reverse causal approach simplifies the identification of efficient mixes of policies with large effects that may not have been known a priori.

Specifically, we model the log of CO_2 emissions (Fig. 1) as a function of log gross domestic product (GDP) and log population and allow for potential breaks in emissions in any country at any point in time that are captured by 'indicators': interactions of country and year fixed effects. Altogether, with an EU-15 sample and 23 time periods, the maximum number of 345 potential treatments exceeds the number of observations. However, countries are treated sparsely so that most indicators are statistically insignificant. We rely on machine learning to remove all but the significant ones (Methods). Those remaining show treatments that significantly reduced country-specific CO_2 emissions relative to a control group conditional on log GDP and log population. It is important to emphasize that these breaks are detected relative to the specified model conditional on the control variables. For example, unconditional visual inspection of Fig. 1 might suggest a break in Greece's CO_2 emissions around 2009. However, the visual 'break' in Greek emissions could be explained by the drop in economic activity due to Greece's sovereign debt crisis. Once we condition on GDP (by including it as a control variable), there is no unexpected change in emissions (and thus no break detected), as emissions were falling in line with GDP.

Having identified a series of breaks, we subsequently attribute the significant indicators to policies and disregard those that show increases for this paper. We construct approximate confidence intervals around an indicator's timing to accommodate for uncertainty. These may be as short as a single year or may span several. Then, we search for policies implemented in these confidence intervals (Methods). In the rare event that an interval incorporates only a single policy intervention, attribution is made with respect to a single policy. Otherwise, attribution is made for a policy mix. Attribution using this approach is no different from arguing that a known intervention is exogenous or as-if random when addressing forward causal questions (discussion in Supplementary Note 1).

Table 1 | Detected breaks, break dates and magnitudes

Country		Model					
		1	2	3	4	5	6
		EU-15	EU-15	EU-15	EU-31	EU-31	EU-31
Significance level in break detection		**5%**	**1%**	**0.1%**	**5%**	**1%**	**0.1%**
Denmark	Effect				−0.080		
	SE				(0.020)		
	Year				2012		
	95% CI				±6		
Finland	Effect	−0.103	−0.123	−0.128	−0.156	−0.171	
	SE	(0.020)	(0.022)	(0.024)	(0.024)	(0.028)	
	Year	2000	2000	2000	2000	2000	
	95% CI	±2	±2	±2	±1	±2	
Germany	Effect	−0.105	−0.131	−0.108	−0.112	−0.112	
	SE	(0.018)	(0.020)	(0.022)	(0.021)	(0.025)	
	Year	2002	2002	2002	2003	2003	
	95% CI	±2	±1	±3	±3	±4	
Ireland	Effect	−0.087		−0.127			
(1st break)	SE	(0.020)		(0.023)			
	Year	2011		2011			
	95% CI	±3		±2			
Ireland	Effect	−0.148	−0.192		−0.247	−0.244	−0.229
(2nd break)	SE	(0.028)	(0.028)		(0.030)	(0.034)	(0.037)
	Year	2015	2015		2015	2015	2015
	95% CI	±1	±1		±0	±1	±1
Luxembourg	Effect	−0.136			−0.108		
(1st break)	SE	(0.024)			(0.031)		
	Year	2007			2007		
	95% CI	±1			±3		
Luxembourg	Effect			−0.214	−0.193	−0.227	−0.262
(2nd break)	SE			(0.031)	(0.030)	(0.035)	(0.038)
	Year			2015	2015	2015	2015
	95% CI			±1	±1	±1	±1
Portugal	Effect				−0.094		
	SE				(0.021)		
	Year				2011		
	95% CI				±4		
Sweden	Effect	−0.095	−0.103	−0.110			
(1st break)	SE	(0.017)	(0.019)	(0.022)			
	Year	2001	2001	2001			
	95% CI	±2	±2	±3			
Sweden	Effect				−0.108	−0.115	
(2nd break)	SE				(0.019)	(0.022)	
	Year				2006	2006	
	95% CI				±3	±4	

This table shows treatment effects on CO_2 emissions, standard errors (SE), the year of the break and its half interval. CI, confidence interval. All treatment effects are statistically significant at the 0.1% level.

The combination of European Union-wide, technological standards but largely diverse national policies across EU member states provides an ideal test bed to learn about effective policy mixes. EU CO_2 efficiency standards for new vehicles have been in place since 1998 and they became mandatory for each EU member state in 2009. In contrast, tax policies, which are the main instruments to achieve national ESR targets, and command-and-control measures vary considerably across member states and time. Prime examples are frequent changes in fuel taxes, the introduction of carbon taxes and road tolls. Moreover, several changes of CO_2-based vehicle taxation

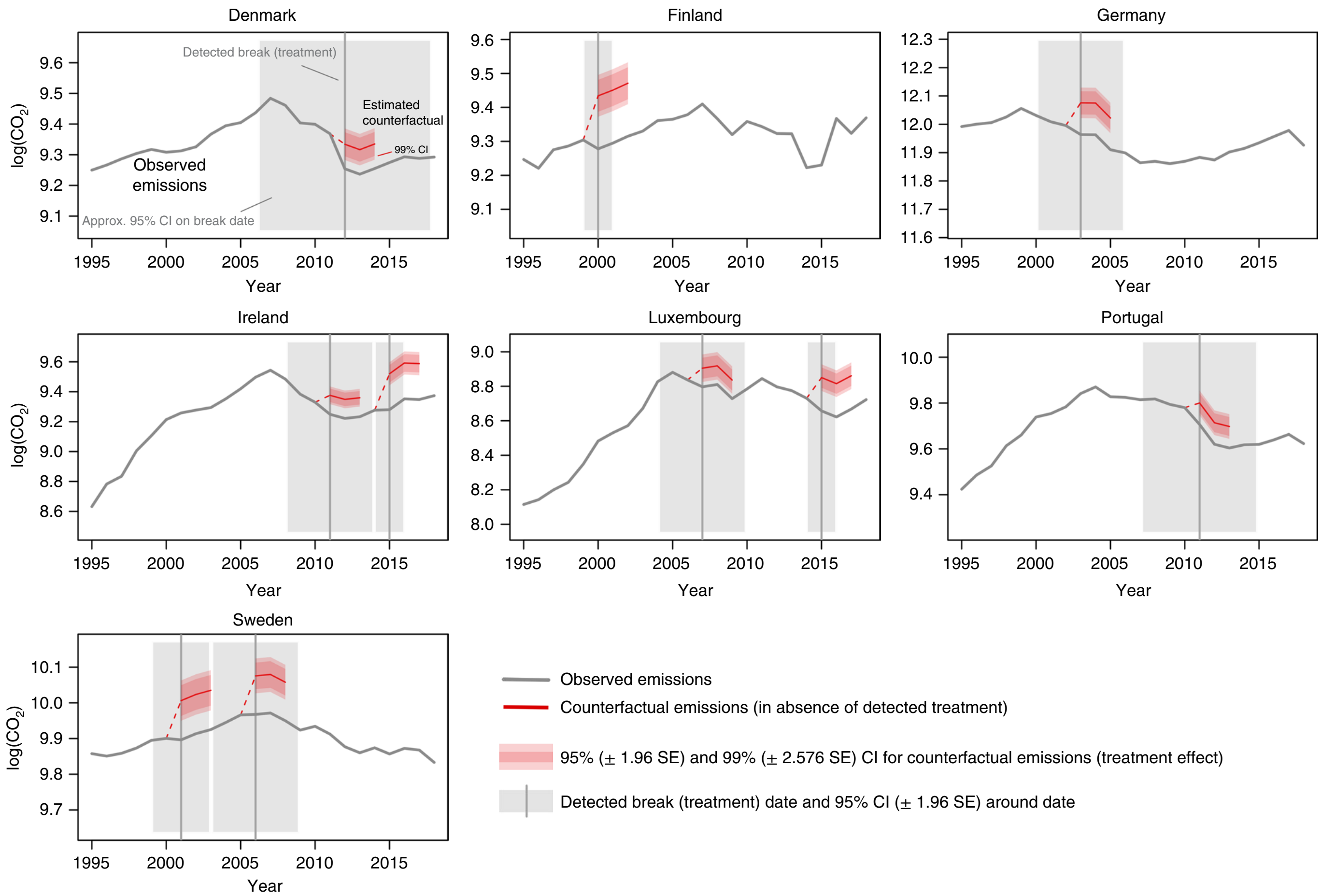

Fig. 3 | Actual and counterfactual road CO_2 emissions. This Figure contrasts the actual emissions (grey line) with their counterfactuals (red) that would have occurred in the absence of treatment (detected as breaks). Counterfactuals are constructed as $\log(CO_2)$ in absence of detected treatment breaks and plotted for three years following the break date. The 95% confidence intervals around the breaks are ± 1.96 standard errors (SE) around their mean estimates; the 99% intervals are ± 2.576 SE wide. The vertical lines indicate detected break dates determined using break detection (Methods). The grey shaded areas indicate an approximate 95% or ± 1.96 SE confidence interval around the timing of the breaks. CI, confidence interval. Approx., approximate.

intended to nudge consumers to buy more fuel efficient vehicles. In addition, the introduction of tax credits and purchase subsidies for electric vehicles and alternative fuels have been popular policies to increase demand for new technologies. Speed limits, biofuel obligations and efficiency labels exemplify varying command-and-control measures across countries and time.

Policy mixes that effectively reduced emissions

We estimate reductions in emissions for the EU-15 members using either a sample of (i) EU-15 members or (ii) EU-31 members, which includes Norway, Iceland, Switzerland and the United Kingdom because they were part of the European Single Market and subject to harmonized regulations. The intuition for a second sample is that it increases the credibility of breaks that are consistently detected across both. We investigate the choice of three target levels of significance for detecting treatments. The target levels of 5%, 1% or 0.1% specify the expected false positive rate of the break detection (Methods). The combination of two samples and three target levels leads to a total of six different models.

Figure 2 shows that only ten of the 345 potential treatments across all countries, years and models are significant. The greater the diameter of a circle indicating a break point, the smaller the level of significance at which it is found in any of the models. The darker its colour, the higher the magnitude of its effect. Table 1 presents detailed results. Altogether, the ten reductions in CO_2 emissions are from seven different countries. Figure 3 shows these relative to the estimated counterfactual (plotted in red for three years following each break date) given by $\log(CO_2)$ in absence of the estimated breaks (estimated as coefficients on the detected break variables). We show the estimated counterfactuals at the estimated break dates (shown as grey vertical lines); however, there is uncertainty around the break dates (shown as grey shading), thus the counterfactuals may visually appear steeper and more sudden than the actual true underlying (albeit unknown) policy effects.

Altogether, we identify six of the ten reductions at a target significance level (and thus expected false positive rate) of 0.1% (Fig. 2). We regard this rate as the most important criterion in break detection to control the level of confidence we require for a policy to be identified. With a rate of 0.1%, we set a very high bar to dispel any lingering doubts that our findings may be driven by spurious false positives. Second, the more models that detect an intervention, the more confident we are of identifying a break. For instance, we find that five of our six models indicate breaks for Finland in 2000, for Germany in 2002/2003 and for Ireland in 2015 (Table 1). We detect that six out of our ten breaks occur in both the EU-15 and the EU-31 sample. Finally, we consider the stability of our estimated effects across models as a third criterion of robustness (Table 1). For instance, for Luxembourg, the coefficients vary between −0.193 and −0.262.

Table 2 | Attribution of detected breaks to policies

Country	Year	Policy
Denmark	2012	2008: Carbon tax increase from 13€ t^{-1} CO_2e to 23€ t^{-1} CO_2e
		2010: 'Green ownership tax' replaces weight-based taxes for light commercial vehicles
		2010: Vehicle tax increase for cars without particle filters
Finland	2000	1996–1999: Carbon tax increases from 2.3€ t^{-1} CO_2e in 1996 to 18€ in 1999
		2001: Car ownership tax base changed from total mass to CO_2 emissions
Germany	2002/2003	1999–2003: 'Ecological Tax Reform' increases motor fuel tax annually by 0.0307€ l^{-1} over five years
		2001: Harmonization of commuter tax deduction between transport modes
		2004: Mandatory fuel efficiency labelling for passenger vehicles
		2005: Road tolls for trucks (originally planned for 2003)
Ireland	2011	2008: Vehicle registration tax base and annual motor tax base shifts from engine size to CO_2 emissions
		2009: Tax incentives for purchase of bicycles for commuting of up to 1,000€
		2009: Electric vehicle subsidy scheme and vehicle registration tax relief
		2010: Introduction of a 15€ t^{-1} CO_2e carbon tax
		2010: Biofuel obligations require blending 4% (6%) biofuels in 2010 (2013)
Ireland	2015	2014: Carbon tax increase to 20€
Luxembourg	2007	2007: Vehicle tax reform based on CO_2 emissions
		2007: Subsidy for purchase of energy efficient vehicles of 750€
		2007–2008: 'Kyoto Cents' law raises fuel tax by 0.02€ l^{-1} for gasoline and 0.025€ l^{-1} for diesel
Luxembourg	2015	2013–2014: Subsidies for electric vehicles and vehicles with < 60 g km^{-1} CO_2
		2015: VAT raise from 15% to 17% increases tax burden of fuelling and buying vehicles
Portugal	2011	2007: Vehicle ownership tax reform based on CO_2 emissions
		2008: Increase of fuel tax by about 0.025€ l^{-1}
		2010: Financial incentives to purchase electric vehicles
		2012: Introduction of nationwide road tolls on motorways and trunk roads
		2015: Introduction of a 5€ t^{-1} CO_2e carbon tax
Sweden	2001	2001–2006: 'Green Tax Shift'
		(i) Carbon tax increase from 40€ in 2000 to 57€ in 2001 to 100€ in 2006
		(ii) Exemptions for biofuels from energy and carbon taxation since 2002
		(iii) Tax benefits for green company cars since 2002
Sweden	2006	2001–2006: 'Green Tax Shift'
		(i) Carbon tax increase from 57€ t^{-1} CO_2e in 2001 to 100€ in 2006
		(ii) Exemptions for biofuels from energy and carbon taxation since 2002
		(iii) Tax benefits for green company cars since 2002
		2005: Pump Act mandates fuel stations to supply biofuel
		2006: Introduction of congestion charges in Stockholm
		2007–2009: Subsidy of up to 1,000€ for eco-friendly vehicles
		2008–2009: Carbon tax increase from 100€ t^{-1} CO_2e in 2006 to 110€ in 2008 to 114€ in 2009

Taking the coefficients with the highest magnitude of relative emissions reductions (in %) and the emissions level (in Mt CO_2) in a given country at the time of the identified break indicates that the breaks we identified between 1995 and 2018 accounted for total emissions reductions of up to 35.9 Mt CO_2. In comparison, the current ESR requires a 30% reduction until 2030 compared with 2005 levels in the sectors that are not subject to EU emissions trading, that is, agriculture, buildings and transport. This target translates into absolute emissions reductions of about 480 Mt CO_2 in the EU-15 member states between 2021 and 2030. If we conservatively assume that each sector contributes in equal measure, the magnitude achieved by past transport policies seem inadequate—even more so when we account for the imminent tightening of the ESR targets to a reduction of 40% relative to 2005 emissions under the proposed revision of the ESR under the 'Fit-For-55' policy package. At the same time, the magnitude of three of the ten detected breaks exceeds 17%, which indicates considerable potential for future reductions.

Post-estimation, we can now attribute effects to their likely causes (Table 2) by matching policies with the break points' confidence intervals. Attribution reveals that many interventions are applied simultaneously often by one legislative package.

For example, we attribute the break point in Luxembourg's emissions around 2007 (99% CI: ± 1 year) to three potential policies: a CO_2-based vehicle tax reform in 2006, a subsidy scheme for fuel efficient cars in 2007 and a 0.02€ per liter fuel tax increase in 2007 ('Kyoto Cents'). The two breakpoints in Sweden occur in 2001 (99% CI: ± 2–3 years) and 2006 (99% CI: ± 3–4 years). These correspond to carbon tax increases implemented in 2001 and tightened annually through 2006. In particular, in 2001, Sweden raised CO_2 taxes from 40€ to 57€ per ton and also introduced subsidies for biofuels and green company cars. Annually increasing CO_2 taxes reached 100€ per ton in 2006. The break in 2006 (±3–4 years) also coincides with the introduction of biofuel mandates, the implementation of road tolls in Stockholm in 2006, the introduction of subsidies for low-emission vehicles in 2007 and further carbon tax increases to 110€ in 2008 and 114€ in 2009, underlining the importance of considering policy mixes rather than individual policies, specifically as multiple detected breaks in a single country could also capture time-varying treatment effects through tightening emissions targets.

We continue by classifying the identified policies. We differentiate between taxes on carbon, fuel or vehicles, road tolls, subsidies and command-and-control measures in Fig. 2. This helps to assess whether certain policies or mixes are superior. We produce four key findings.

First, we link all detected emissions reductions to at least one tax policy that increases the cost of driving; we link seven cases to carbon, four cases to fuel taxes and three cases to road tolls. Second, we link eight of the ten emissions breaks to policies that combine taxes that increase the cost of driving with reforms that emphasize CO_2-based vehicle taxes (six cases) or subsidy schemes for low-emissions vehicles (six cases). For instance, we attribute the ten to 15% emissions reduction in Finland in 2000 to a combination of a carbon tax increase and switching to CO_2-based vehicle taxes. Vehicle taxes and subsidies provide incentives to switch to more fuel efficient or zero emissions vehicles, in particular, if consumers either systematically underestimate or discount future savings from increased efficiency. Third, we link the breaks with the highest level of confidence and the greatest magnitude of effect (Finland 2000, Germany 2002/2003, Luxembourg 2015, Ireland 2015) to increases in existing but moderate carbon or fuel taxes.

Fourth, our finding that command-and-control measures relate only to three emissions breaks (mandatory efficiency labels in Germany and biofuel obligation schemes in Ireland and Sweden) potentially indicates that they either play a minor role in reducing CO_2 emissions at the national level or that governments did not use them extensively. However, we caution against over-interpreting this finding because key command-and-control measures, such as efficiency standards for new vehicles, are implemented at the EU level. We can detect measures only at the national level, which might be of limited impact. Moreover, our search for potential policy measures relies on databases that hardly include any public transport policies.

Finally, we note our approach's limitations. One concern is that agnostic break detection runs the risk of not detecting real but less effective treatments. To address this concern, we use higher target levels of significance (that is, expected false positive rates) that allow identification of smaller and therefore more potential treatments (Fig. 2). As a further robustness check, we also searched our policy databases for carbon, fuel and road tax interventions that we do not detect (Table 3). Figure 4 compares all actually implemented carbon tax changes to the ones we detected. Overall, we detect all but two. The lack of evidence for any emission break in France despite its 2014 carbon tax introduction may be best explained by the fact that the initial tax of 7€ was offset by an equivalent reduction in the existing energy consumption tax[22]. Similarly, the lack of finding any effects for the 2011 carbon tax increase in Finland may be because of simultaneous reductions in the tax on engine power for cars and trucks that might have weakened its effect[23]. We do not find any major undetected toll increases except for one in Austria in 2004 and the introduction of a vignette system in the United Kingdom in 2014. However, Table 3 shows that we do find a number of undetected (sometimes transitory) changes in fuel taxes that exhibit a wide range of magnitudes. Potentially relevant but undetected fuel tax increases occur in Austria, Belgium, Italy, the Netherlands, Spain and the United Kingdom. The emissions effect of these tax changes is likely too small to be identified by our approach. We draw two conclusions from these robustness checks. First, we are more likely to detect large breaks and cannot detect effective policies that yield small reductions. Thus, our estimates for the effect sizes of our detected policy mixes provide a lower-bound estimate if countries in the control group also experienced smaller emissions reductions that we do not detect. Given the magnitude and urgency of the climate crisis and the ambitious EU climate targets, we believe a focus on interventions with large-scale effects is justified. Second, we caution against generalizing our results and using them as benchmark estimates for particular policy instruments or policy mixes. The provision of such benchmark estimates is an important policy question for future research that can be tackled by combining our proposed reverse causal approach with the standard forward causal approach (Supplementary Note 1 for a more detailed discussion).

Table 3 | Undetected carbon, fuel or road pricing policies

Country	Year	Undetected Policies
Austria	2004	Introduction of electronic network-wide road toll system for trucks (which increased costs compared with the previous vignette system)
Austria	2008	Increase of fuel tax by about $0.03€\,l^{-1}$
Austria	2012	Increase of fuel tax by about $0.04€\,l^{-1}$
Belgium	2006	Increase of fuel tax by about $0.08€\,l^{-1}$
Belgium	2010	Increase of fuel tax by about $0.02€\,l^{-1}$
Finland	2010–2012	Increase of carbon tax from about 20€ to about $60€\,t^{-1}\,CO_2e$
France	2014	Introduction of a $7€\,t^{-1}\,CO_2e$ carbon tax
Greece	2008–2012	Gradual increase of fuel tax from about $0.33€\,l^{-1}$ in 2008 to $0.67€\,l^{-1}$ in 2012
Italy	2006	Increase of fuel tax by about $0.02€\,l^{-1}$
Italy	2012	Increase of fuel tax by about $0.14€\,l^{-1}$
Netherlands	2005	Increase of fuel tax by about $0.04€\,l^{-1}$ (and subsequently annual increases by about $0.01–0.02€\,l^{-1}$)
Spain	2010	Increase of fuel tax by about $0.065€\,l^{-1}$
United Kingdom	2012	Increase of fuel tax by about $0.06€\,l^{-1}$ (back to tax level before 2010)
United Kingdom	2014	Introduction of road toll vignette system for trucks

Data from ACEA Tax Guide[43], CESifo DICE Report[44], World Bank's Carbon Pricing Dashboard[45] and country-specific sources specified therein.

Given that we detect breaks relative to a specified model conditional on selected control variables, a related concern is that our model may be mis-specified and might lead to the detection of spurious breaks. However, Supplementary Tables 8–11 in Supplementary Note 3 show that our findings are generally robust to various alternative baseline model specifications (including new controls such as the share of urban population, nonlinear functional forms and linear country-specific time trends), especially based on our preferred EU-15 control group. There are two notable exceptions: (i) the break in Portugal, which already had weak support in our main Table 1 and, as expected with country-specific time trends, (ii) including time trends absorbs the two breaks in Sweden that indicate time-varying treatment effects from the Swedish climate policy package. In addition, a specification test suggested by Oster[24] shows that our results are robust with respect to omitted variable bias (Supplementary Table 12 in Supplementary Note 3).

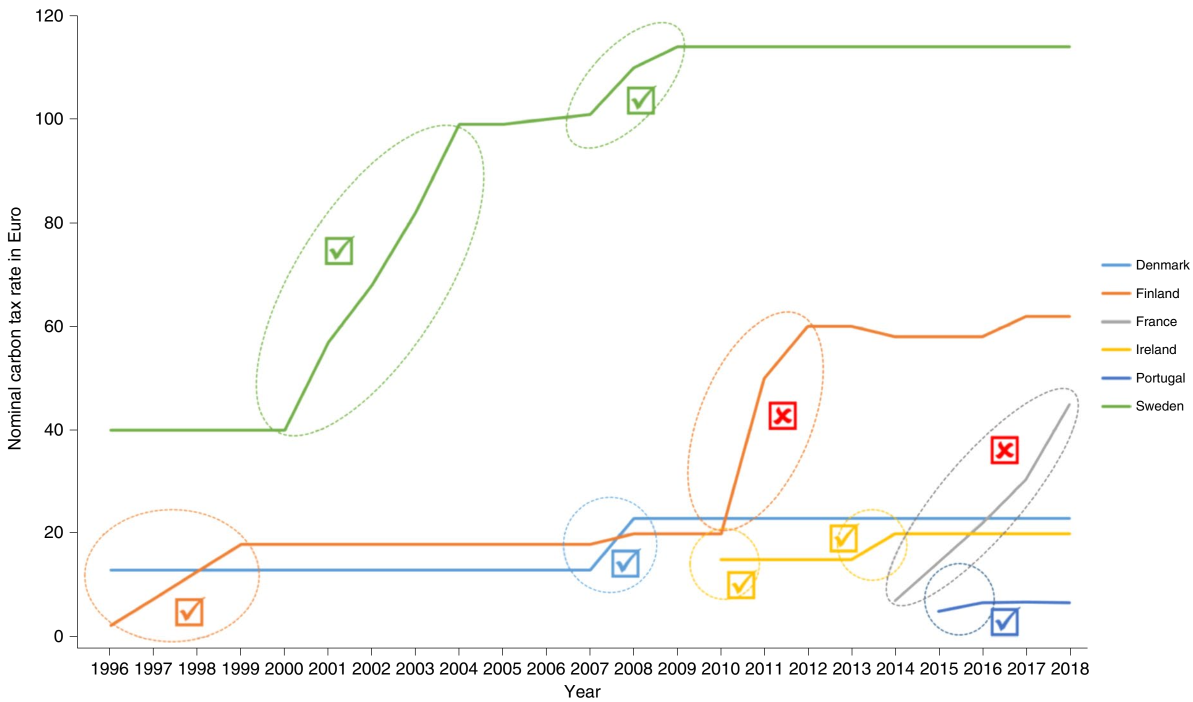

Fig. 4 | Overview of implemented and detected carbon tax changes. Nominal carbon tax rates in €$t^{-1}CO_2$ between 1995 and 2018 for all EU-15 countries with a carbon pricing scheme. We encircled relevant changes in carbon tax rates. A tick indicates that we detect the tax changes, while a cross indicates that we do not. Data from World Bank's Carbon Pricing Dashboard[45] and country-specific sources specified therein.

Another concern is that the estimated effects may not be generalizable because (i) we estimate country-specific effects for each intervention and (ii) we are more likely to detect large breaks. To partially address the latter concern, we report bias-adjusted coefficients that do not change the interpretation of our results in Supplementary Table 6 in Supplementary Note 3. The former concern may be addressed by averaging over all identified and similar treatments to approximate the average treatment effect (in the spirit of ref. [25]). But we do not seek to provide benchmark estimates for particular policy instruments here.

A final concern is that national policies may also affect neighbouring countries. In particular, a fuel or carbon tax increase may cause fuel tourism in private consumers in the border region or cause firms to reroute their trucks for refuelling. To evaluate the potential bias from such spillovers, we exclude the two major transit countries with a low fuel tax regime, Austria and Luxembourg, from our sample to estimate our final model. Supplementary Table 7 in Supplementary Note 3 shows that we obtain very similar results in this restricted sample. Moreover, these concerns also apply to forward causal analyses.

Conclusion

In this study, we propose a complementary approach to ex post policy evaluation. Instead of estimating the effect of a single, known cause on emissions, we seek to identify multiple, unknown causes of an emissions effect. As policymakers implement ever more climate policy packages to meet their obligations under the Paris Agreement or their own net-zero emissions targets, we believe our approach is policy relevant because it enables drawing systematic inference on the effectiveness of such policy *mixes*. We demonstrate this for the EU transport sector, which is a key bottleneck that impedes the European Union's progress to achieve climate neutrality by 2050.

Our results show that relatively few policy interventions effectively curbed CO_2 emissions in road transport. We identify ten successful interventions with emissions reductions between 8% and 26% or 35.9 Mt CO_2 between 1995 and 2018 and attribute all detected emissions reductions to policy mixes that comprise at least one tax policy intervention that increases the cost of driving. The fact that we detect nearly all carbon price interventions indicates that carbon pricing may be a critical element of effective policy packages. In addition, we attribute the vast majority of emissions reductions to policy mixes that combine carbon, fuel or road-use taxes with additional vehicle taxes or subsidies. The most successful examples of such combinations are policies implemented in Finland in 2000 (−17%) Sweden in 2001 (−11%), Ireland in 2011 (−13%) and Luxembourg in 2015 (−26%). Carbon, fuel or road-use taxes provide incentives to reduce mileage, yet they may not ensure that consumers invest in energy efficient vehicles if consumers are myopic. This effect is known as the energy efficiency gap[26]. Vehicle taxes and subsidies can address myopic consumers and provide incentives to adopt more fuel efficient vehicles. However, they suffer from the rebound effect that describes the unintended side effect that more efficient vehicles cost less to drive and, therefore, encourage additional mileage[27]. Our findings thus provide suggestive evidence that the combinations of policies that simultaneously address the energy efficiency gap and rebound effects are particularly effective. To check the robustness of this evidence based on country-specific effect estimates, future research may combine our proposed reverse causal approach with the standard forward causal approach to provide more systematic benchmark estimates for the effectiveness of particular policy mixes. Our findings are broadly in line with studies that use more structural modelling to evaluate policy mixes[28,29] and the literature suggesting that tax policies can address rebound effects[30,31]. Finally, we also show that the greatest emissions reductions occur when policymakers increase existing but moderate carbon or fuel taxes. This suggests that commitment to staggered, anticipated and permanent tax increases over time may be a strong determinant of emissions reductions.

Altogether, the ambitious country-specific emissions reduction targets under the EU effort sharing regulation require timely action.

We identified policy mixes with emissions reductions on a magnitude that matches the reduction requirements under the net-zero emissions target for seven EU countries. If policymakers in these countries focused on the policy mixes that have been effective in the past, we should expect stronger reductions in road transport emissions. Although policies are context-specific, we yet believe that policymakers in other EU countries may also learn from these successful interventions.

Methods

Data. The data for road transport CO_2 emissions is from section 1A3b of the Emissions Database for Global Atmospheric Research (EDGAR) v5.0 (ref. [32]). We retrieved GDP and data on population sizes from World Bank[33,34]. The dependent variable is the natural logarithm of CO_2 emissions log(CO_2). log(GDP), log(GDP)^2 and log(population) enter the model as control variables. Supplementary Table 4 in Supplementary Note 3 shows that our findings are robust, if we restrict CO_2 emissions to passenger vehicles only (that is, section 1A3bi).

We analysed emissions break in the EU-15 member states (Austria, Belgium, Germany, Denmark, Spain, Finland, France, United Kingdom, Ireland, Italy, Luxembourg, Netherlands, Greece, Portugal and Sweden) because they are subject to largely identical EU regulations (with some minor differences in implementation), in general, and specifically with respect to the European Single Market over our sample period from 1996 to 2018. We disregard years before 1996 due to major historic dissimilarities between these countries. For the control group, we also consider a broader sample of all 27 EU member states and the three European Free Trade Association states and the United Kingdom (EU-15, Croatia, Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Lithuania, Latvia, Malta, Poland, Romania, Slovakia, Slovenia, Switzerland, Iceland and Norway). Supplementary Tables 1–3 in Supplementary Note 2 provide summary statistics.

Empirical approach. We identify effective policy interventions by detecting structural breaks in TWFE panel models of CO_2 emissions. Detected breaks identify heterogeneous treatment effects without prior knowledge on treatment assignment or timing. A standard approach to analyse policy effectiveness, often interpreted as difference-in-differences when treatment effects are homogeneous—for example, ref. [13]—is to model emissions using a TWFE estimator as a function of control variables and a binary variable that denotes the interaction of 'treated' countries that are subject to particular policies and the post-treatment period. Such 'known' binary policy variables in a TWFE panel are equivalent to step shifts in the individual fixed effects of the treated countries (more detailed discussion in Supplementary Note 1).

Using the equivalence between step shifts in the unit-specific intercept (that is, fixed effect) and known treatments, we use an alternative approach to evaluate reverse causal questions regarding policy interventions. Rather than exclusively evaluating known interventions while disregarding unknown but effective policies, we estimate a TWFE panel in search of potential structural breaks (step shifts) in the unit-specific intercepts. Once a break has been identified, it can be interpreted as a treatment for the relevant country. We then attempt to attribute the break to a policy that affected the treated country around the detected time. Thus, rather than assessing effects of causes, our approach provides a data-driven method to first identify breaks which can, in a second step, be attributed to policy interventions. Pretis and Schwarz[17] provide a detailed discussion of this modelling approach that was first introduced by Pretis[35].

We formulate the detection of structural breaks as a problem of variable selection similar to ref. [36] but extended the approach to the panel setting, where we saturate a TWFE panel model with a full set of step shifts denoting potential treatment of every country at every point in time. We then apply variable selection methods from machine learning that allow for more candidate variables than observations to identify breaks without prior knowledge of their existence. We saturate a TWFE regression with a full set of break variables (step shifts) denoting potential treatment of each unit at every time period, nesting any specific treatment as a special case. In a balanced panel of N countries and T time periods this adds $N(T-1)$ potential break variables to be selected over. Therefore, we start with a full set of step functions with coefficients $\tau_{j,s}$:

$$\log\left(\mathrm{CO}_2\right)_{i,t} = \alpha_i + \phi_t + \sum_{j=1}^{N}\sum_{s=2}^{T} \tau_{j,s} 1_{\{i=j,t\geq s\}} + x'\beta + \epsilon_{i,t} \tag{1}$$

where α_i and ϕ_t denote individual and time fixed effects, $x_{i,t}$ is a vector of control variables that includes log(GDP), log(GDP)^2 and log(population size). The population treatment coefficients $\tau_{j,s}$ are sparse with coefficients of zero for all but the treated countries. This operationalizes the notion of ref. [11] that reverse causal questions require variables '*that might not yet even be in our model*.' The target of model selection is then to remove all but the relevant break variables so that in a final sparse model, the selected breaks correspond to the true underlying, and potentially unknown treatments. Let $\widehat{\mathrm{Tr}}$ denote the set of detected treated countries, with associated detected treatment times $\hat{T}_j$ for each treated country $j \in \widehat{\mathrm{Tr}}$. Then the resulting sparse model is:

$$\widehat{\log(\mathrm{CO}_2)}_{i,t} = \hat{\alpha}_i + \hat{\phi}_t + \sum_{j\in\widehat{Tr}}\sum_{s\in\hat{T}_j} \hat{\tau}_{j,s} 1_{\{i=j,t\geq s\}} + x'\hat{\beta} \tag{2}$$

Coefficients $\hat{\tau}_{j,s}$ correspond to estimates of heterogeneous treatment effects for the detected treated countries. For example, we may detect a break for Sweden in 2006 (Swe $\in$ Tr, $T_{\mathrm{Swe}} = 2006$), where the associated estimated coefficient $\hat{\tau}_{\mathrm{Swe},2006}$ on the break variable captures the country-specific treatment effect. The estimated treatment effects in the final retained model (2) can be interpreted as heterogeneous treatment effects estimated using interactions of the unit-fixed effects with treatment time for each treated unit as in ref. [25], thus also addressing recent concerns about imposing homogeneous treatment effects in panels with staggered adoption (for example, ref. [37]). The main difference relative to the specification in ref. [25] is that in our application, each treated cohort consists of a single country.

We resort to machine learning to move from the general model (1) that embeds all possible treatment dates for all countries to the sparse model (2). A large set of potential selection algorithms are available. To carefully control the false positive rate of detected breaks, we apply the block search algorithm 'gets'[38] using the 'getspanel' update in ref. [18], which forms part of the general-to-specific family of model selection. Alternatives include shrinkage-based methods such as the LASSO and variants thereof, though these do not target the false positive rate (refs. [39–41]). Supplementary Note 1 provides a more detailed discussion.

The main calibration parameter of 'gets' is the target level of significance γ_c which controls the expected false positive rate of retained breaks and is defined as the number of non-zero treatment coefficients relative to all possible treatment coefficients. Their asymptotic properties are explored in ref. [42], who show that in the absence of breaks and accounting for multiple testing, the false positive rate converges to the chosen nominal level of significance of selection γ_c. If there are no true treatment breaks present, then the proportion of spuriously detected breaks converges to the chosen level of significance. For instance, with $\gamma_c = 0.01$, the expected false positive rate is 1% and we expect $0.01 \times N(T-1)$ spuriously retained breaks. We consider γ_c equal to 0.05, 0.01 and 0.001 in our models of CO_2 emissions to assess the robustness of our results. Supplementary Table 5 in Supplementary Note 3 shows that our findings are robust to using cluster-robust standard errors.

Attribution. Our attribution strategy to match policy interventions to the year intervals for which we detect break points involved two primary databases and various supplementary data sources.

First, we searched for interventions in two main databases: (i) the IEA's Policies and Measures Database that provides information on past, existing or planned climate and energy policies. Data is collected from governments, international organizations and IEA analyses, and governments can review the provided information periodically. (ii) The National Communications to the United Nations Framework Convention on Climate Change secretariat that our sample countries are required to submit regularly.

Second, to corroborate the information gained from the IEA and United Nations Framework Convention on Climate Change documents and to double check for any policies these two sources omit, we collected additional information from the European Automobile Manufacturers' Association's Annual Tax Guide that provides detailed information on fuel, vehicle and road tax schedules and subsidy programmes, the World Bank's Carbon Pricing Dashboard that provides detailed information on carbon prices and the Climate Change Laws of the World database of the Grantham Research Institute. In a few cases, we also conducted specific searches on Google.

Data availability

All publicly available data analysed in this study are available from the corresponding author upon request and are also available from online repository Zenodo (https://doi.org/10.5281/zenodo.6768563).

Code availability

The code required to replicate our study is available from the corresponding author upon request and is also available from online repository Zenodo (https://doi.org/10.5281/zenodo.6768563).

Received: 28 February 2022; Accepted: 12 July 2022;
Published online: 22 August 2022

References

1. *Net Zero by 2050—A Roadmap for the Global Energy Sector* (IEA, 2021).
2. Axsen, J., Plötz, P. & Wolinetz, M. Crafting strong, integrated policy mixes for deep CO_2 mitigation in road transport. *Nat. Clim.Change* **10**, 809–818 (2020).

3. *Planning for Net Zero: Assessing the Draft National Energy and Climate Plans* (Ecologic & Climact, 2019).
4. *Trends and Projections in Europe 2021* (European Environment Agency, 2021).
5. Graf, A., Graichen, J., Matthes, F. C., Gores, S. & Fallasch, F. *How to Raise Europe's Climate Ambitions for 2030* (Agora Energiewende and Öko-Institut e.V., 2019).
6. Transport and Energy *Why Increasing Ambition under the ESR is Unavoidable* (2021).
7. Grant, D., Bergstrand, K. & Running, K. Effectiveness of US state policies in reducing CO_2 emissions from power plants. *Nat. Clim. Change* **4**, 977–982 (2014).
8. Martin, G. & Saikawa, E. Effectiveness of state climate and energy policies in reducing power-sector CO_2 emissions. *Nat. Clim. Change* **7**, 912–919 (2017).
9. Andersson, J. J. Carbon taxes and CO_2 emissions: Sweden as a case study. *Am. Econ. J.* **11**, 1–30 (2019).
10. Bayer, P. & Aklin, M. The European Union emissions trading system reduced CO_2 emissions despite low prices. *Proc. Natl Acad. Sci. USA* **117**, 8804–8812 (2020).
11. Gelman, A. & Imbens, G. *Why Ask Why? Forward Causal Inference and Reverse Causal Questions* (National Bureau of Economic Research, 2013).
12. Eskander, S. M. & Fankhauser, S. Reduction in greenhouse gas emissions from national climate legislation. *Nat. Clim. Change* **10**, 750–756 (2020).
13. Lin, B. & Li, X. The effect of carbon tax on per capita CO_2 emissions. *Energy Policy* **39**, 5137–5146 (2011).
14. Klemetsen, M., Rosendahl, K. E. & Jakobsen, A. L. The impacts of the EU ETS on Norwegian plant's environmental and economic performance. *Clim. Change Econ* **11**, 2050006 (2020).
15. Colmer, J., Martin, R., Muûls, M. & Wagner, U. J. *Does Pricing Carbon Mitigate Climate Change? Firm-Level Evidence from the European Union Emissions Trading Scheme*, discussion paper 1728 (Center for Economic Performance, 2020).
16. Rafaty, R., Dolphin, G. & Pretis, F. *Carbon Pricing and the Elasticity of CO_2 Emissions*, working Paper No. 140 (Institute for New Economic Thinking, 2020).
17. Pretis, F. & Schwarz, M. Discovering what mattered: Answering reverse causal questions by detecting unknown treatment assignment and timing as breaks in panel models. *Preprint at SSRN* https://doi.org/10.2139/ssrn.4022745 (2022).
18. Schwarz, M. & Pretis, F. getspanel. GitHub repository (2021), https://github.com/moritzpschwarz/getspanel/
19. Estrada, F., Perron, P. & Martínez-López, B. Statistically derived contributions of diverse human influences to twentieth-century temperature changes. *Nat. Geosci.* **6**, 1050–1055 (2013).
20. Hendry, D. F. et al. *First In, First Out: Econometric Modelling of UK Annual CO_2 Emissions, 1860–2017* (Economics Group, Nuffield College, Univ. of Oxford, 2020).
21. Piehl, A. M., Cooper, S. J., Braga, A. A. & Kennedy, D. M. Testing for structural breaks in the evaluation of programs. *Rev. Econ. Stat.* **85**, 550–558 (2003).
22. Schubert, K. Carbon taxation: The French experience, 2014–2019. *Coalition of Finance Ministers for Climate Action Workshop on Carbon Taxation*, 2019.
23. *Report of the Working Group on Energy Taxation Reform: A Proposal for Implementing the Intentions and Goals of the Government Programme and for Further Development of Energy Taxation* (Finnish Ministry of Finance, 2021).
24. Oster, E. Unobservable selection and coefficient stability: theory and evidence. *J. Bus. Econ. Stat.* **37**, 187–204 (2019).
25. Wooldridge, J. Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. Preprint at *SSRN* https://doi.org/10.2139/ssrn.3906345 (2021).
26. Gillingham, K. T., Houde, S. & van Benthem, A. A. Consumer myopia in vehicle purchases: evidence from a natural experiment. *Am. Econ. J.* **13**, 207–38 (2021).
27. Gillingham, K., Kotchen, M. J., Rapson, D. S. & Wagner, G. The rebound effect is overplayed. *Nature* **493**, 475–476 (2013).
28. Ravigné, E., Ghersi, F. & Nadaud, F. Is a fair energy transition possible? Evidence from the French low-carbon strategy. *Ecol. Econ.* **196**, 107397 (2022).
29. Landis, F., Rausch, S., Kosch, M. & Böhringer, C. Efficient and equitable policy design: taxing energy use or promoting energy savings?. *Energy J.* **40**, 73–104 (2019).
30. Vivanco, D. F., Kemp, R. & van der Voet, E. How to deal with the rebound effect? A policy-oriented approach. *Energy Policy* **94**, 114–125 (2016).
31. Freire-González, J. & Ho, M. S. Policy strategies to tackle rebound effects: a comparative analysis. *Ecol. Econ.* **193**, 107332 (2022).
32. Crippa, M. et al. *Population, Total Fossil CO_2 and GHG Emissions of All World Countries—2019 Report* (Publications Office of the European Union, 2019).
33. GDP (constant 2010 US$). *World Bank Open Data* (2020), https://data.worldbank.org/indicator/NY.GDP.MKTP.KD?view=chart
34. Population, total. *World Bank Open Data*, https://data.worldbank.org/indicator/SP.POP.TOTL?view=chart (2020).
35. Pretis, F. Does a carbon tax reduce CO_2 emissions? Evidence from British Columbia. *Environmental and Resource Economics* (2022), https://doi.org/10.1007/s10640-022-00679-w
36. Castle, J. L., Doornik, J. A., Hendry, D. F. & Pretis, F. Detecting location shifts during model selection by step-indicator saturation. *Econometrics* **3**, 240–264 (2015).
37. Goodman-Bacon, A. Difference-in-differences with variation in treatment timing. *J. Econometrics* **225**, 254–277 (2021).
38. Pretis, F., Reade, J. & Sucarrat, G. Automated general-to-specific (gets) regression modeling and indicator saturation methods for the detection of outliers and structural breaks. *J. Stat. Softw.* **86**(3), 1–44 (2018).
39. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **58**, 267–288 (1996).
40. Zou, H. The adaptive lasso and its oracle properties. *J. Am. Stat. Assoc.* **101**, 1418–1429 (2006).
41. Okui, R. & Wang, W. Heterogeneous structural breaks in panel data models. *J. Econometrics* **220**, 447–473 (2021).
42. Nielsen, B. & Qian, M. Asymptotic properties of the gauge of step-indicator saturation. Working Paper, University of Oxford (2018).
43. *ACEA Tax Guide* (ACEA) (2022).
44. Rumscheidt, S. Road user charging in the European Union. *CESifo DICE Rep.* **12**, 54–57 (2014).
45. *Carbon Pricing Dashboard* (World Bank, 2022); https://carbonpricingdashboard.worldbank.org/

Acknowledgements

We thank O. Edenhofer, A.B. Martinez, R. Tol and the participants at the EC2 Conference 2021, the Federal Reserve Virtual Seminar on Climate Economics and the Climate Econometrics Seminar for valuable feedback and suggestions. F.P. and M.S. gratefully acknowledge funding from the Clarendon Fund and the Robertson Foundation. F.P. is also grateful to funding from Social Sciences and Humanities Research Council of Canada (SSHRC). The views expressed here are those of the authors and not necessarily those of the Ministry of Finance or the Austrian government.

Author contributions

F.P., L.N., M.S. and N.K. designed the analysis. F.P. and M.S. wrote the core programme code. L.N. collected the data. L.N. and N.K. conducted most of the analyses. All authors interpreted results and designed figures. N.R. and N.K. wrote the manuscript with contributions from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41560-022-01095-6.

Correspondence and requests for materials should be addressed to Nicolas Koch.

Peer review information *Nature Energy* thanks Patrick Bayer, Edgar Hertwich and Md. Saniul Alam for their contribution to the peer review of this work.