STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2020-2021

The Swedish Outdoor Smoking Ban of 2019: An Application of Synthetic Control and Matrix Completion Methods in Policy Evaluation

Petter Berg (41636) and Julia Blomstergren (23907)

Abstract: Smoking bans constitute an important part of the tobacco control policies used globally to improve public health. In July 2019, Sweden implemented new regulation which banned smoking in several outdoor areas nationwide, including outdoor serving areas of e.g. restaurants and bars, whose scope is unique in an international context. This paper aims to estimate the short-term impact of this ban on cigarette consumption using retail scanner data in a cross-country context. Employing Synthetic Control methods as well as the recently proposed Matrix Completion with Nuclear Norm Minimization estimator we find no evidence for the existence of an effect of the ban, nor do we find any visual signs of trend breaks in related outcomes such as sales of a major substitute tobacco product or proxies for quit attempts. Our findings contribute to the literature on tobacco control measures by providing estimates based on sales data instead of self-reports, and to the policy evaluation literature in general by providing an example of how recent developments in methodological tools can be used in application. Finally, we provide decision makers deliberating on whether to follow Sweden's example in legislation with scientific evidence on the topic.

Keywords: Tobacco Control Policy, Smoking Bans, Counterfactual Prediction, Synthetic Controls, Matrix Completion, Policy Evaluation

JEL: I12, K32, I18, C33

Supervisor:	Abhijeet Singh
Date submitted:	May 15, 2021
Date examined:	May 25, 2021
Discussants:	Elias Höckerfelt & Johanna Ohlin
Examiner:	Magnus Johannesson

Acknowledgements

We would like to thank our supervisor Abhijeet Singh for his valuable input during the process of writing this thesis. We also extend our gratitude to Julius Andersson and Chloe Nibourel for their helpful feedback on our application of the synthetic control method. Further, we would like to thank Anna Ekström at Nielsen for her helpfulness in making this study possible, as well as Joanna Stjernschantz Forsberg and Anna Fält at Sluta-Röka-Linjen for providing us with data.

Contents

1 Introduction **1**

2 Background **3**
2.1 An overview of tobacco control policy 4
2.2 The Swedish smoking ban of 2019 5

3 Literature review **6**
3.1 The effect of smoking bans 6
3.2 Previous studies of the Swedish smoking ban of 2019 10

4 Data **10**
4.1 Data on cigarette sales 10
4.2 Aligning the periods of the cigarette sales data 12
4.3 Data on covariates and country characteristics 12

5 Empirical analysis **13**
5.1 Visual analysis of Sweden's cigarette sales 14
5.2 Characteristics of Sweden and the donor pool 16
5.2.1 Smoking consumption and demographics 16
5.2.2 Tobacco control policies 19
5.2.3 Climate and weather 21
5.3 The Synthetic Control Method 23
5.3.1 Conceptual framework 23
5.3.2 Theory 24
5.3.3 Analysis and results 26
5.3.4 Relaxing the country weight constraints 32
5.4 Matrix Completion with Nuclear Norm Minimization 35
5.4.1 Conceptual framework 35
5.4.2 Theory 36
5.4.3 Analysis and results 38

6 Robustness checks and extensions **43**
6.1 Transformation of the outcome variable 43
6.2 Trends in snus sales 49
6.3 Trends in efforts to quit smoking 50
6.4 A note on statistical power 51

7 Discussion **53**

8 Conclusion **58**

Appendices **65**

1 Introduction

Smoking is a dying phenomenon, at least if you ask Sweden's policy makers. Due to the catastrophic impacts of tobacco smoking on public health, policy makers worldwide have taken up the fight against smoking and seen smoking rates fall drastically during the latest decades—a development that has been underpinned by a vast amount of research on the effectiveness of various tobacco control policies. Yet, Sweden's government is among those who have declared that they will not settle for anything less than a smoke-free society, which they strive to achieve by 2025. On July 1, 2019, the most recent step in this direction was taken when a smoking ban was implemented that extended the country's smoking restrictions to cover several outdoor areas.

This new law banned smoking in outdoor serving areas of bars, restaurants and clubs, at public transportation stops, playgrounds, sports facilities as well as at the entrances of all facilities to which the public has access, such as stores. The stated intent of this ban was twofold. First, it was expected to reduce secondhand smoke exposure, thereby increasing access to public spaces for people who are particularly sensitive to smoke (Public Health Agency of Sweden 2019). As a second objective, the government expected smoking per se to decline, with reference to the strong scientific evidence on the inhibitory effects of indoor smoking bans on cigarette consumption (Prop. 2017/18:156). However, in the bill, it is recognized that the scientific evidence on *outdoor* smoking bans, such as the one proposed, is not unequivocal. The purpose of this paper is to contribute to this relatively thin body of research and investigate whether the smoking ban, which we will hereafter refer to as the smoking ban of 2019, has been successful in its aim to reduce cigarette consumption. In doing so, we fail to find any effect of the smoking ban over the studied period of July 2019 through January 2020.

In an international comparison, the Swedish smoking ban of 2019 is the first of its kind in terms of the wide range of outdoor areas covered in combination with the nation-wide implementation. Similar partial bans have been implemented previously in several countries, but only at the local or regional level.[1] In general, outdoor smoking bans are much less common than indoor smoking bans, and those that exist often apply only to parks, school zones or beaches. Hence, most previous studies on the effect of smoking bans on cigarette consumption have focused on indoor bans in e.g. workplaces or other public spaces which are much more widely adopted. Early research on such bans found large effects in the order of 5–25% (Levy, Chaloupka, and Gitchell 2004), while more recent evidence is mixed (Anger, Kvasnicka, and Siedler 2011; Boes, Marti, and Maclean 2015; Jones et al. 2015). The literature on outdoor smoking bans and their impact on cigarette consumption is thin, amounting to two studies: Kennedy (2010), which finds an effect on self-reported smoking, and Ko (2020) in which findings with respect to the existence of an effect are ambiguous. Our study of the Swedish smoking ban of 2019 will therefore contribute to the literature on smoking control policies by examining a weaker

[1] Examples include the state of Victoria in Australia (State Government of Victoria 2015) and four provinces in Canada (Azagba 2015).

extension of the indoor smoking bans that are implemented in most Western countries, that is simultaneously stronger than the more common outdoor bans solely confined to parks, school zones and the like.

One of the main problems with studying the Swedish smoking ban of 2019, and any other event that took place around the same time, is that the unprecedented shock of Covid-19 hit the world in early 2020. There are few if any social phenomena that have remained unaffected by the pandemic. Further, since countries' strategies to combat the pandemic have differed vastly, with Sweden's approach being rather unique in a global context, the magnitude of the pandemic's impact can be expected to vary across countries. In econometric terms, this means that the Covid-19 pandemic represents a shock with heterogeneous effects across countries, rendering the effects of the smoking ban on a one to a few years horizon fundamentally unidentifiable. What is more, identification of longer term effects will necessarily rely on the assumption that the pandemic-induced changes to smoking behavior are not permanent regime shifts, which may prove quite difficult to either prove or disqualify. The time frame of our study is therefore restricted to the short run, and more precisely the seven months between the date at which the smoking ban came into force until late January 2020 just before the pandemic began.

Since the smoking ban of 2019 was implemented at a national level, we use cross-country analysis methods to study its effects. However, a common issue in cross-country analyses is that the control units are too different from the treated unit to provide reasonable counterfactuals, so that e.g. assumptions of parallel trends in methods such as Difference-in-Differences (DID) are not sufficiently well grounded. Recent years have seen a surge in econometric literature trying to develop methods that relax these assumptions in various ways, with a prominent example being the synthetic control method (hereafter referred to as the SC method) developed in a series of papers by Alberto Abadie and co-authors (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010; Abadie, Diamond, and Hainmueller 2015). This method takes a data-driven approach to constructing a counterfactual to the treated unit as a linear combination of the outcomes of the control units. More recently, Athey et al. (2021) proposed an estimator based on nuclear norm minimization, closely related to the matrix completion literature, in order to achieve a low-rank approximation of the outcome matrix that can subsequently be used to predict counterfactual outcomes for the treated unit. In order to investigate the impact of the smoking ban of 2019, this paper takes the SC method as a methodological starting point and extends the analysis by employing the Matrix Completion with Nuclear Norm Minimization estimator (hereafter referred to as the MC method) proposed in Athey et al. (2021).

As our measure of cigarette consumption, we use monthly retail scanner data on cigarette sales that we have acquired from Nielsen, which is a global consumer research firm with a particular focus on providing high quality market measurements and insights. Cigarette sales were recorded on a monthly basis over a three-year period covering nine countries.[2] For our purposes, the main benefit of using retail scanner data over other

[2]The countries are Austria, Estonia, Ireland, Latvia, the Netherlands, Norway, UK, USA and Sweden.

potential sources, such as tax records, is that it provides a highly accurate measurement of the *timing* of sales. Since we are investigating a short-term effect on consumer behavior, this aspect of the data is of paramount importance. Tax data, on the other hand, are based on reports that do not necessarily coincide with the time of the actual sale but rather depends on the country-specific rules regarding when the tax liability arises. Further, retail scanner data capture revealed consumption as opposed to the stated consumption of e.g. surveys that are commonly used in the literature on smoking, which allows us to avoid issues related to desirability bias (perhaps particularly pressing when related to activities commonly perceived as "bad habits") or survey respondents' imperfect recall of past activities (Fisher and Katz 2000; Brzozowski, Crossley, and Winter 2017).

On basis of our analysis, we find that the Swedish smoking ban of 2019 had no substantial impact on cigarette sales during the seven post-treatment months studied. This result is robust over all specifications and estimators employed in this paper. In particular, we attain point estimates that are very close to zero or even positive, i.e. the opposite direction of what we would expect, which are consistently not extreme enough to be distinguishable from the typical prediction error of around 5% of the pre-treatment mean. Based on these findings, we arrive at the conclusion that any effect of the smoking ban greater than 5% of the pre-treatment mean of cigarette sales in Sweden can be ruled out, which corresponds to approximately 0.075 cigarette packs or 1.5 cigarette sticks per capita and month. Our findings are important for Swedish authorities from a policy evaluation perspective, for authorities in countries considering the implementation of similar outdoor smoking bans, and for other scholars as it extends our knowledge regarding a topic on which few previous studies have been carried out.

This paper is structured as follows. In the following section, we describe the background of tobacco control policy globally and the context in which the Swedish smoking ban of 2019 is situated. Further, we describe in detail the design of, reasoning behind and implementation of the ban. In Section 3, we investigate previous literature on the effects of indoor and outdoor smoking bans. Section 4 provides an exposition of the data that we use in analysis of the Swedish ban. In Section 5, we present the methodological details as well as the main analysis and results of the paper. In Section 6, we provide robustness checks of our main analysis and briefly investigate trends in the sales of snus, a widely used tobacco product in Sweden that smokers might have substituted cigarettes for, and efforts to quit smoking. Further, we briefly discuss the statistical power of our study. This is followed by a discussion of the results in Section 7, and Section 8 concludes.

2 Background

In this section, we begin with elaborating on the status of and motivation behind tobacco control policies in general. After doing so, we go into the specifics of the Swedish smoking ban of 2019 both in terms of legislative design and the expectations of an effect on part of policy makers.

2.1 An overview of tobacco control policy

Tobacco-related mortality and morbidity present significant threats to public health globally, with the World Health Organization (WHO) calling the situation a global tobacco epidemic. Specifically, consumption of tobacco products and exposure to secondhand tobacco smoke increase the risk of several of the most deadly non-communicable diseases, such as cardiovascular diseases, cancers and respiratory diseases (WHO 2019). Nonetheless, some 10 to 50 percent of the adult population in Western countries consume tobacco products regularly, of which cigarettes are by far the most common one. The negative tobacco-related health effects incur large costs both on the individuals affected, in terms of increased risk of premature mortality, foregone earnings and reduced life satisfaction due to diseases, and society through for example health care costs and dropouts from the labor force (Chaloupka and Warner 2000). In Sweden, it is estimated that 18,000 deaths annually and 8% of the total disease burden can be ascribed to tobacco consumption or exposure (Socialstyrelsen 2014). On a global scale, the WHO estimates that the annual premature mortality from tobacco-related diseases amounts to eight million people, of which 10% are non-smokers exposed to secondhand tobacco smoke (WHO 2020). Because of this, both national governments and international organizations alike are working to combat the consumption of tobacco through policy interventions. The objectives of such tobacco control policies can be thought of as twofold: i) protecting non-smokers and ii) decreasing tobacco consumption by lowering the consumption per tobacco user and/or reducing its prevalence through discouraging non-smokers from starting and encouraging smokers to quit (Hopkins et al. 2001).

To facilitate the establishment of effective tobacco control policies, the WHO developed the WHO Framework Convention on Tobacco Control which was signed by 168 countries, including Sweden, between 2003 and 2004. This treaty was among the earliest international initiatives coordinating tobacco control policies, and has been highly impactful. It serves as a guideline for countries, mapping out which areas tobacco control policies should target and what types of legislation should be used, based on what research has found to be effective (WHO 2005). In the European area, the EU has also been an important agent for tobacco control policies by for example standardizing the labelling and packaging of tobacco products, issuing recommendations and in other ways working to support and advance the member countries' policies (European Commission 2015a; European Commission 2017). Due to the high degree of international collaboration on and standardization of policies, tobacco legislation is fairly similar across countries, in particular in a Western context. Examples of these similarities include that all countries relevant in the context of this study use specific tobacco taxes, have banned or strongly restricted smoking in public indoor spaces, and demand warnings to cover large parts of tobacco products' packaging (Joossens, Feliu, and Fernandez 2020). The Swedish smoking ban of 2019 becomes particularly interesting against this background, seeing that it is aligned with the tobacco control policy strategies recommended by the WHO and EU, yet not a kind of regulation that has been widely adopted internationally. While outdoor

smoking bans exist in many countries, they are mostly limited to campuses, beaches, parks or areas in conjunction with hospitals and schools, and many of them are implemented on a regional or local scale, especially the more extensive ones. Hence, the Swedish ban stands out with its inclusion of a wide range of outdoor areas and national coverage. If the Swedish policy proves effective, however, it is possible that other countries will follow by implementing similar bans.

2.2 The Swedish smoking ban of 2019

The new Swedish tobacco law, of which the outdoor ban under study in this paper is part, was passed in 2018, regulating the sale and use of tobacco products and similar goods such as e-cigarettes (SFS 2018:2088). This law replaced previous legislation that was divided into two pieces; *Tobakslagen* [The Tobacco Law] (SFS 1993:581) and *Lagen om elektroniska cigaretter och påfyllningsbehållare* [The Law on E-Cigarettes and Replacement Containers] (SFS 2017:425). The new law mandates that, in addition to previously regulated areas such as indoor public spaces, smoking of tobacco products and similar products such as e-cigarettes and herbal products is banned in a number of outdoor environments: outdoor areas in conjunction with public transportation, outdoor serving areas of for example restaurants, bars and cafés, sports facilities, publicly accessible playgrounds as well as entrances to publicly accessible premises, such as shops and malls.

The new smoking restrictions came into effect on July 1, 2019, and the motivation behind it is twofold (Prop. 2017/18:156). First, banning smoking in the aforementioned outdoor spaces is expected to reduce exposure to secondhand smoke, which is beneficial from a public health perspective. It is also argued on grounds of fairness that people who are particularly sensitive to or dislike smoke exposure should have a greater access to these public spaces, which the ban would accomplish according to legislators.

The smoking ban is, however, not only expected to relocate smoking to areas where fewer are exposed to the secondhand smoke, but also to reduce cigarette consumption per se. In particular, the ban is embedded in the broader context of a vision of a smoke-free Sweden by 2025, that has been endorsed by the government. The expectation of an effect on smoking is primarily motivated by the findings of a public inquiry performed by the Public Health Agency of Sweden (Public Health Agency of Sweden 2014). In their report, the Public Health Agency of Sweden states that there is strong scientific support for the existence of a negative effect on smoking of indoor smoking bans. They have also evaluated the effectiveness of a number of extensive local outdoor smoking bans that have been implemented in for example Canada, Australia and the United States. Based on this, they conclude that there is reason to believe that the proposed outdoor smoking ban will be effective. Hence, the prior on part of legislators is that the proposed ban would reduce smoking, although it is also noted that scientific evidence on outdoor bans is scarce.

The smoking ban mandates that the person that owns or disposes of the public space in question is responsible for enforcing the ban and informing the public about it by way

of signs (SFS 2018:2088). If a person is violating the ban and does not respond to verbal actions, the owner holds the right to dismiss the person with help from police or security guards. Municipalities are responsible for the surveillance of the ban's enforcement. Shortly after the ban came into force, news outlets reported on high compliance in outdoor serving areas of restaurants and bars, which essentially turned smoke-free overnight, but less so at unsupervised public spaces such as bus stops (Nyberg 2019; Ritzén 2019). The view that the ban has been successful in making outdoor serving areas smoke-free is shared by The Hotel and Restaurant Workers' Union (HRF), that has long advocated for such legislation (Niia et al. 2009; Cullberg 2019). A possible explanation to the reported high degree of compliance in outdoor serving areas is that enforcing the ban is aligned with the incentives of staff to avoid exposure to secondhand smoke.

A final important implication of the smoking ban of 2019 is that it is no longer possible to consume alcohol and cigarettes in direct conjunction in for example restaurants and bars. The smoking ban mandates that in order to smoke, a customer must leave the physical boundaries of the facility by a sufficiently large distance so as to ensure that other customers are not affected, which also excludes the possibility of smoking at its entrance (SFS 2018:2088). At the same time, however, the customer has to leave any purchased alcohol within the premises of the serving area since it is strictly forbidden in Sweden for served alcohol to leave the premises at which it was sold (SFS 2010:1622).

3 Literature review

In this section, we review previous literature on the effect of smoking bans on cigarette consumption. To the best of our knowledge, there are very few previous studies on the effect of outdoor smoking bans on smoking prevalence and/or intensity. Therefore, we complement our review of these studies with an overview of the large body of literature on indoor smoking bans. Although such interventions are stronger than the outdoor smoking ban in study, the mechanisms are arguably similar enough to make it an imperfect but still informative comparison.

3.1 The effect of smoking bans

Smoking bans constitute an important class of tobacco control policy tools. They address both of the objectives of tobacco control policies that have been previously discussed: protecting non-smokers from secondhand smoke by creating smoke-free environments and, in theory, reducing cigarette consumption by increasing the marginal cost and decreasing the marginal benefits of smoking (Levy, Chaloupka, and Gitchell 2004). In relation to the first of these objectives, previous literature has documented large positive health effects of smoking bans in various environments, both indoors and outdoors (Azagba 2015; Frazer et al. 2016). However, this strain of research falls outside of the scope of this paper, as our research question revolves around the success of the smoking ban of 2019 in relation to the latter objective of reducing smoking.

We begin with outlining the theoretical argument for smoking bans' effect on smoking behavior, that is often set in a cost-benefit analysis framework where larger (smaller) costs (benefits) are predicted to decrease consumption. The most straightforward impact of smoking bans on smokers' costs is that the time and effort costs of smoking that would otherwise have taken place in the regulated environments increase, as smokers must relocate to areas in which smoking is allowed. This effect may however be partially offset by more intensive smoking in expectation of spending time in smoke-free environments, especially for heavily addicted smokers making short visits to such environments (Chandra and Chaloupka 2003). Smoking bans also prevent social smoking, i.e. smoking motivated by the social context, in the environments that are declared smoke-free. As such, smoking bans applying to places where social smoking is common such as bars and nightclubs may reduce the benefits of smoking that would otherwise have taken place at those locations unless smokers can relocate to an environment equally suitable for social smoking (Shiffman et al. 2012; Vuolo, Kelly, and Kadowaki 2016). Odermatt and Stutzer (2015) further suggests that smoking bans can act as self-control devices for smokers attempting to cease or reduce their cigarette consumption. Specifically, smokers will experience less temptations in the regulated areas and be better equipped to resist their impulses to smoke, since the time from the impulse to when the cigarette can be lit is prolonged. Around 70% of smokers typically report that they would like to quit but have been unable to, suggesting that self-control mechanisms may have an important role to play (Babb et al. 2017).[3] In addition to the instant effects discussed above, smoking bans are expected to have long-term impacts on smoking behavior. By denormalizing smoking, smoking bans are expected to diminish the role of smoking in the culture, decreasing the appeal of picking up smoking and possibly increasing the cost of smoking due to it being a socially stigmatized activity (Stuber, Galea, and Link 2008; Frazer et al. 2016).

On the empirical side, the impact of smoking bans on smoking behavior has been studied in many different ways in the literature. These studies vary in several aspects, such as the time-frame studied (short- or long-term), the measure of smoking behavior used (intensity or prevalence), and the type of data used (sales or survey data). A majority of the studies focus on restrictions to smoking in indoor public spaces such as workplaces and/or the hospitality sector—restaurants, bars and similar establishments—since this is the most common type of smoking ban. Banning indoor smoking in public venues should however be seen as a stronger treatment than the one being studied in this paper, since the induced marginal increase in the cost of smoking is presumably larger, and since most people spend a larger fraction of their time in public indoor spaces than in public outdoor spaces. As previously mentioned, this view is shared by the legislators behind the Swedish smoking ban of 2019 (Prop. 2017/18:156). Therefore, the indoor smoking bans that we review may only serve as imperfect comparisons to the Swedish ban. Levy, Chaloupka, and Gitchell (2004) provides an early survey of the literature on indoor smoking bans and finds that clean air laws are important for reducing smoking, especially

[3]Although this number may be inflated due to desirability bias, it is high enough to indicate that a substantial share would remain in absence of bias.

in combination with informational campaigns. For example, U.S. studies have found correlation between the comprehensiveness of clean air laws and cigarette consumption on the state level, with differences in the per capita consumption of 5–20%, and similar results are found in European countries. A study by Bauer et al. (2005) also finds that employees in workplaces adopting smoke-free policies between 1993 and 2001 were almost twice as likely to have quit smoking by the end of the period, compared to those whose workplaces did not forbid smoking, while also reducing the cigarette consumption for continuing smokers by almost 4 cigarettes daily, amounting to almost 6 packs a month. However, in two more recent reviews of the literature, Wilson et al. (2012) and Frazer et al. (2016) both find inconclusive evidence with respect to the impact of indoor smoking bans on smoking behavior, with some reviewed studies finding immediate effects and/or changes in the trend and other studies finding no effect. As such, while policy makers have great confidence in smoking bans' ability to reduce smoking, the evidence from some of the research is rather mixed. However, it should also be mentioned that many of the reviewed studies employ methods that provide limited support for causal inference, such as before-after analyses, and should therefore be viewed only as indicative.

A number of more recent studies stemming from the empirical health economics literature have investigated the effect of indoor smoking bans in public venues using DID approaches. Anger, Kvasnicka, and Siedler (2011) looks explicitly at the short-term impacts (less than 1 year) of such bans in Germany, and finds no effect in the general population but a strong decline in prevalence and intensity of smoking among individuals who frequent bars and restaurants. This result is complemented by other studies, such as Vuolo, Kelly, and Kadowaki (2016), which find that smoking bans are primarily effective in reducing cigarette consumption among social smoking young adults, which is a group that also appears to be more sensitive to price changes (Gallet and List 2003). Another similar paper investigates the gradual implementation of smoking bans in public venues in Switzerland and finds a 6.3% reduction in smoking prevalence relative to the control group that appears after one year of treatment (Boes, Marti, and Maclean 2015). However, by using a binary indicator for smoking this paper only captures effects along the extensive margin of cigarette consumption and not the potential additional impacts on the intensity of consumption. Finally, Jones et al. (2015) uses the same DID methodology in a UK setting, exploiting the fact that England and Scotland introduced indoor smoking bans in public venues at different points in time during the late 2000's. The authors find suggestive but not statistically significant evidence of a short-term (less than 1 year) reduction in cigarette consumption following the ban.

Next, we turn to studies of outdoor smoking restrictions, that may serve as more fitting comparisons to the Swedish smoking ban of 2019 than the indoor bans reviewed above. The number of studies investigating the effect of outdoor smoking bans on smoking rates or intensity is very small, and most previous literature has focused on measuring the compliance with such bans (see e.g. Azagba 2015). In particular, and to the best of our knowledge, there are only two previous studies that investigate the impact on cigarette consumption. The first of these, Ko (2020), studies a South Korean outdoor

smoking ban applying to parks, bus stops and school zones. Exploiting the gradual roll-out of the outdoor smoking restrictions, Ko (2020) finds an increase in the probability of making a smoking cessation attempt of 16% using survey panel data. The effect seems to arise immediately in conjunction with the implementation of the ban and is persistent over multiple subsequent years. Despite this, the study finds no effect on smoking rates and intensity. As noted by Ko, effect sizes of smoking bans are likely contingent on the restrictions applying to the kinds of areas where smokers usually consume their cigarettes; a ban on smoking in parks is likely to be ineffective if people do not wish to smoke there, no matter what policies are in place. Compared to the ban that Ko studies, the Swedish 2019 smoking ban applies to outdoor environments in which smoking is arguably more desirable, especially the outdoor serving areas of restaurants and nightclubs and around entrances to public spaces. Hence, the Swedish ban is arguably a stronger treatment which may lead to a larger effect size than the one found in Ko (2020).

In the second study on outdoor smoking restrictions, Kennedy (2010) evaluates a smoking ban that was implemented in the Canadian city of Woodstock (pop. 35,480), applying to the same outdoor environments as the Swedish smoking ban—outdoor serving areas, public transportation stops, playgrounds, sports facilities and entrances to public spaces—as well as parks. As such, this ban is comparable to the Swedish ban in terms the areas affected, although one should always be careful in extrapolating results that are estimated on a small population that has likely self-selected into treatment. Nevertheless, one year after implementation Kennedy finds that 15–26% of smokers reported that they are more likely to quit smoking because of the ban, and 30–42% reduced their consumption of cigarettes in response to it. Although such self-reported data should be interpreted carefully due to social desirability bias, the findings are indicative of the ban having effect both on smokers' intent to quit and actual consumption.

We conclude this section by providing a summary of the literature discussed above, and examining the Swedish 2019 smoking ban in the context of it. First off, early studies on indoor smoking bans often found rather sizeable effects on smoking prevalence, intensity and quitting rates. However, in the more recent economics literature, the evidence is more mixed. In particular, short-term effects on the aggregate vary from as much as a 6.3% reduction in smoking *prevalence* (Boes, Marti, and Maclean 2015), to essentially null (Anger, Kvasnicka, and Siedler 2011; Jones et al. 2015). In our view, contributing factors to the findings being mixed may be that it has been difficult to disentangle effects of various tobacco control policies that have been implemented within short periods of time, and that effects may be dependent on contextual factors such as the social norms in relation to smoking and degree of enforcement. If we turn to outdoor smoking bans' effects on cigarette consumption, the evidence is scarce, with Ko (2020) finding effects on cessation attempts but not consumption overall, and Kennedy (2010) finding large effects on the intent to quit and self-reported consumption. However, one caveat with all of these studies is that they rely on self-reported measures of smoking prevalence and/or intensity, that may be prone to bias.

Given this context, the Swedish smoking ban of 2019 is essentially uncharted territory

as it constitutes a stronger treatment than the national outdoor smoking ban studied by Ko (2020) and implemented on a different scale compared to the local outdoor ban studied by Kennedy (2010), while arguably being weaker than the previous wave of smoke-free indoor policies implemented during the 2000's and early 2010's. Hence, by investigating this unexplored middle ground using data based on actual sales rather than survey data, our study contributes to the literature on tobacco control policies in general while also providing important insights to countries aiming to reduce cigarette consumption that are deliberating on whether to follow Sweden's example.

3.2 Previous studies of the Swedish smoking ban of 2019

To the best of our knowledge, there have been no attempts thus far to robustly estimate the impact of the smoking ban of 2019 on smoking behavior. Zetterqvist and Ramstedt (2019) briefly analyzes the ban in a recent report published by The Swedish Council for Information on Alcohol and Other Drugs (CAN) by comparing the share of smokers and consumption of cigarettes per capita in the first half of 2019, before the smoking ban came into effect, with the second half of the year based on survey data. Such a before-and-after analysis without controls does however not account for potential sources of bias arising from seasonal patterns in smoking not being evenly distributed around July 1, when the ban came into effect, or the existence of a general time trend in cigarette consumption more broadly. Therefore, we argue that the analysis by Zetterqvist and Ramstedt provides inadequate support for any causal claim regarding the effect of the smoking ban of 2019.

4 Data

In this section, we describe the data used in this study. We focus on the cigarette sales data, that constitute the main outcome studied in this paper, but also briefly discuss data sources for the various predictors and covariates used.

4.1 Data on cigarette sales

The main data source used in this paper is retail scanner data on cigarette sales, obtained from Nielsen Corporation (henceforth Nielsen). Nielsen is a global measurement and data analytics company that, among many other things, collects high quality data through collaborations with retailers in a wide range of industries. In particular, these data are collected via checkout scanners such that not only the quantity but also the exact timing of sales is recorded. Before making data accessible for corporate partners or researchers, such as ourselves, Nielsen applies various control procedures in order to ensure that the data are of high quality (Nielsen Corporation 2021). In most of the analysis, we express the cigarette sales data in number of cigarette packs sold per capita, which we construct by complementing the Nielsen data with population data from the World Bank (2021b).

The cigarette sales data are aggregated on a monthly level, with 12 or 13 observations per year depending on the mode of aggregation, for 9 countries: Austria, Estonia, Ireland, Latvia, the Netherlands, Norway, Sweden, the United Kingdom and the United States. These countries were selected by our contact person at Nielsen on basis of availability from a list of suggestions provided by ourselves consisting of countries deemed sufficiently similar to Sweden in order to be useful for analysis.[4] We also made several attempts to acquire data on additional countries after the initial response, but that was not possible on part of Nielsen. Austria was excluded from the analysis since a major tobacco control policy came into force in November 2019, such that our final data set consists of 8 countries in which no major tobacco control policies were implemented during our study period. The series begin in October 2017 and end in October 2020. However, for matters of analysis we only consider periods before the breakout of the Covid-19 pandemic and therefore truncate the series at the end of January 2020. The data were obtained in late October 2020 and since Nielsen only stores data for 36 months, it was not possible to access data prior to October 2017. In the series that consist of 13 observation per year, the running time variable is four-week periods whereas in the series that contain 12 observations per year it follows calendar months.

As the data are obtained through Nielsen's retail network, the cigarette sales series do not cover the entirety of sales within each country. Using data on annual cigarette sales reported by official sources based on tax records and published by the European Commission (European Commission 2020), we estimate that Nielsen's market coverage varies in the range of 68% to 91% between the countries in our sample.[5] These estimated market coverage rates are presented in Table 1. In Sweden, we estimate the market coverage to around 80%. As a result of Nielsen's data collection methods, this does not represent a random sample of vendors since it covers the sales at the retail chains well but generally leaves out small, independent shops. Our assessment is that the coverage rate is high enough to ensure that we capture the important variation in cigarette sales in our analysis. Moreover, we assess any potential bias stemming from the data measurement structure to be of limited importance. Even if we, for example, were to suspect that the effect of the smoking ban under study would mainly go through sales at small shops, the majority of such small shops (e.g. Pressbyrån or 7-Eleven) are not independent but rather parts of retail chains and are therefore included in the measured data.

[4]This list included Sweden, Norway, Denmark, Finland, UK, Ireland, France, the Netherlands, Belgium, Austria, Switzerland, Poland, Iceland, Estonia, Latvia, Lithuania, USA and Canada. However, it was also specified that data on any additional developed country not on this list would potentially be of interest too.

[5]We calculate these estimates as a fraction with the sum of the country's monthly cigarette sales in 2019 in the numerator and the corresponding sum based on annual data reported by official sources in the denominator. The calculated market coverage rates are only indicative and will not be used in analysis, since the timing of measurement differs between the data sources, and since the annual tax data series can contain artificial shocks caused by stocking behavior in expectation of tobacco control policies implemented with the turn of the year.

Table 1
Estimated market coverage rates of Nielsen by country

	Norway	UK	Latvia	Sweden	USA	Netherlands	Ireland	Estonia
Coverage (%)	91	91	88	79	79	77	70	68

Note: Market coverage rates are calculated as the total cigarette sales in 2019 reported by Nielsen as a percentage of the cigarette sales figures reported by tax authorities in the respective countries in the same year. Sources: Authors' rendering of data from Nielsen and European Commission 2020.

4.2 Aligning the periods of the cigarette sales data

For two of the countries in our data set, Latvia and Estonia, Nielsen have aggregated the cigarette sales data on a calendar month basis, while the rest come in 4 week sums. This presents a problem for analysis, as we need the data to be on the same frequency. Since we do not have the luxury of being able to simply exclude these countries, due to the already rather limited number of control countries, we chose to transform the calendar month data into 4 week sums by interpolation. In particular, we perform an unweighted split of the calendar months into days, and then aggregate them into periods of 28 days (4 weeks). This corresponds to a linear interpolation method since the constructed periods are weighted averages of two adjacent calendar months. Although other, more complex interpolation methods can be considered, we believe that this approach is sufficiently robust for our purposes.

4.3 Data on covariates and country characteristics

The use of covariates play an important role in the SC method as well as the MC estimator. In this section, we provide information on the data sources of these covariates, while a more thorough discussion on the motivation for including them will follow in the respective sections of the empirical analysis. We also provide references for the data sources of the country characteristics presented in section 5.2.

Data on annual cigarette sales between 2013 and 2018, used in several contexts in the paper, have been collected from European Commission (2020) and are based on tax records from the respective countries reported by official sources. We also use temperature data in two ways, collected from two different sources due to differences in availability. First, we compute long-term (2006–2016) monthly average temperatures in order to capture the average temperature differences between summer and winter in the various countries, and the data for this are collected from World Bank (2021a). Second, we use monthly temperature data over the actual period of our outcome variable, capturing the average monthly temperatures between late 2017 and early 2020 by country, and these data are retrieved from Berkeley Earth (2021). Cigarette tax rate data between 2017 and 2019 for all countries except Norway and United States are retrieved from European Commission (2021), while the equivalent measure for Norway and Unites States comes from WHO (2021). Further, data on the share of occasional smokers, as well as prevalence of daily and occasional smoking by gender, in 2018 are retrieved from WHO

(2021), that builds on various representative national surveys. However, this data set does not contain smoking prevalence by age groups, which we have instead collected from the European Health Interview Survey (EHIS) conducted in 2014 and published by European Commission (2015b) for all countries except for the United States, for which we have accessed the equivalent data from Hu et al. (2015). Unfortunately, the EHIS is only performed every fifth year, meaning that the 2014 data is the most recent data that we can use.

5 Empirical analysis

In this section of the paper, we conduct the empirical analysis of the 2019 smoking ban in Sweden. The outcome variable of interest is cigarette sales per capita, measured in packs (one pack contains roughly 20 cigarette sticks), in a given country and time period. In later analysis, this measure is transformed into percentage deviations from the pre-treatment mean, and the reasons for doing so are detailed in section 5.2. All data used in the analysis are aggregated over time periods of four weeks (28 days), with the first and last time periods ending in November 5, 2017, and January 26, 2020, respectively. With the treatment date being July 1, 2019, the last untreated time period ends on June 16, 2019, followed by a partly treated period that ends on July 14, 2019.[6]

Before going into the analysis itself, we briefly describe the model selection process behind it. Since the smoking ban of 2019 was enforced in all of Sweden at the same time, it is not possible to conduct a robust, within-country analysis of the policy. Hence, we turn to the literature on comparative research methods in a panel data setting. The canonical DID approach is a widely used research tool in the economic literature, but is predicated on the existence of a suitable control unit whose pre-treatment trend in the outcome variable is similar to that of the treated unit. In settings where such a unit does not exist, alternative methods can be employed. This is the focus of the synthetic control literature in which a counterfactual is constructed as a weighted average of multiple control units that better matches the outcome of the treated series than any control unit taken alone (Abadie, Diamond, and Hainmueller 2010; Abadie, Diamond, and Hainmueller 2015; Abadie 2020). In the basic version of the estimator, the weights are restricted to be non-negative and sum to unity in order to avoid extrapolation. However, other scholars have proposed variations that relax these assumptions (see e.g. Doudchenko and Imbens 2016). Further, a recent paper by Athey et al. (2021) provides a synthesis on these methodological advancements and develops a new estimator based on matrix completion techniques that, under various restrictions, nests both the DID estimator and the SC method while also, in some settings, outperforming them both.

Our main analysis is based on the aforementioned literature and consists of three parts. First, we apply the original SC method as proposed in Abadie, Diamond, and

[6]The reason for the first post-treatment period being only partly treated is that the period of our data does not coincide with the turn of the calendar month.

Hainmueller (2010). Next, we turn to an extension of this method where the constraints on the control unit weights are relaxed in order to allow for a slight degree of extrapolation, similar to the approach in Doudchenko and Imbens (2016). Finally, we apply the MC method developed in Athey et al. (2021) that can be seen as the most recent development in the literature to date.

This part of the paper is structured as follows. In the following section, we perform a visual analysis of Sweden's cigarette sales over the sample period. Section 5.2 investigates the characteristics of Sweden and the control units available to us, both in terms of the outcome variable and several other features that may influence cigarette consumption such as smoker demographics and climate. In Section 5.3, we describe the SC method in detail and provide results and placebo tests, both from the original formulation as well as the one in which constraints are relaxed. Section 5.4 follows the same structure but for the MC method.

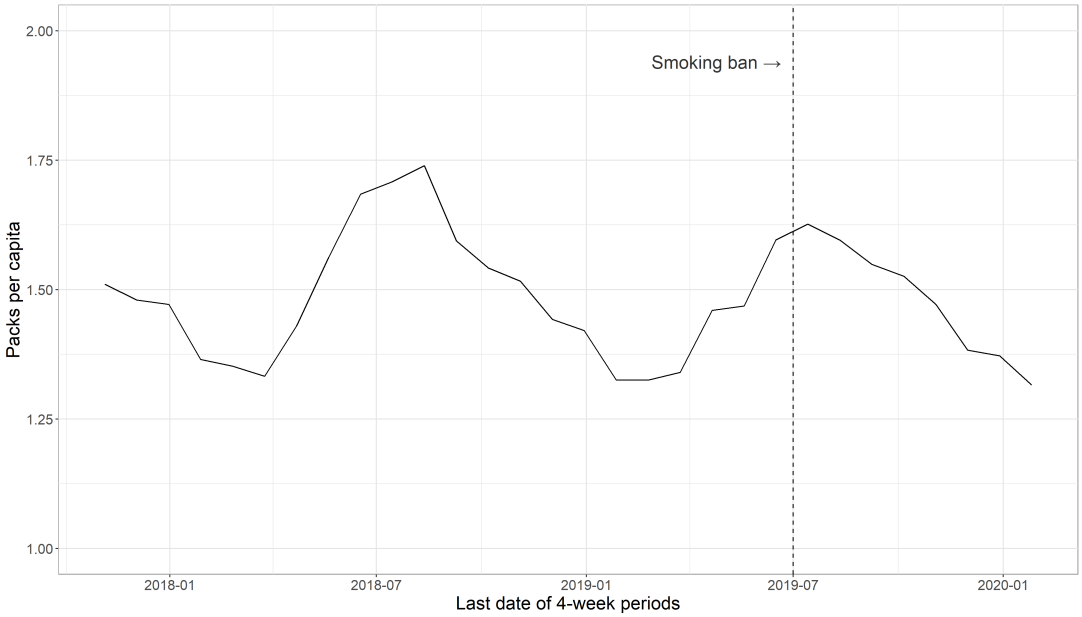
5.1 Visual analysis of Sweden's cigarette sales

It it informative to begin by inspecting the series of interest, i.e. Sweden's cigarette sales, to see whether there are any signs of an interruption or change of slope in the raw data at the time of the implementation of the smoking ban of 2019. While such changes may be driven by unobserved confounders, it gives motivation for further analysis and employing methods that can control for such biases.

Figure 1 shows the cigarette packs sales per capita per 4-week period over our sample period in Sweden. The series clearly exhibits a lower peak in the summer of 2019 compared to the one preceding it. In fact, the upward trend during the spring of 2019 is abruptly broken just after the treatment date, reversing to a downward trend afterwards. In contrast, the positive trend in the spring/summer of 2018 seemed to persist until around August, only thereafter reverting back down during the fall. Of course, there may be many explanations to this that do not involve the smoking ban. For example, if Sweden would be on a long-term negative trend in cigarette consumption, we would expect the peak in 2018 to be lower than that of 2019 in the absence of shocks. However, the time-span of our sales data from Nielsen, stretching back only to the end of 2017, does not allow us to discern the direction or even existence of such a long-term trend.

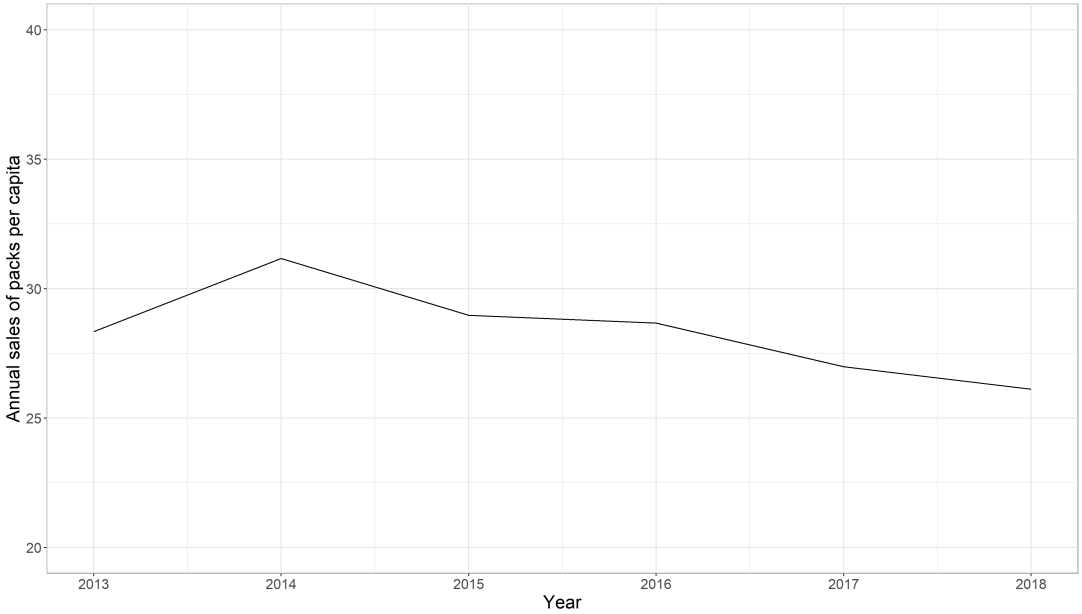
We can, however, investigate the trend by looking at data from tax records that cover a longer time period than the data obtained from Nielsen. Figure 2 shows the annual number of cigarette packs per capita sold in Sweden between 2013 and 2018. The data exhibit a clear downward trend; over this period, the mean growth rate from one year to the next amounted to -1.4%. If we look at the monthly data again, the mean packs per capita sold during the summer months of June through August fell by about 6% between 2018 and 2019, while the sales in the winter months was almost constant, resulting in a total change over the entire years of 2018 and 2019 of -3.6%. This could very well be explained by the general downward trend in annual cigarette sales observed in Figure 2 being driven by continuous reductions in summertime smoking. However, it could just as

Figure 1
Monthly time series of cigarette packs sold per capita in Sweden



Source: Authors' rendering of data from Nielsen.

Figure 2
Annual time series of cigarette packs sold per capita in Sweden



Source: Authors' rendering of data from European Commission (2020).

well be that the rather large reduction during the summer months is a result of the 2019 smoking ban, while the almost constant sales over the two winters is an anomaly. Finally, since long-term trends are almost never perfectly linear in practise we simply cannot tell whether the change between 2018 and 2019 would have been at, above or below the mean of -1.4% holding everything else constant.

In the end, making any causal statement on basis of the visual inspection of a single time series is difficult due to the absence of a proper control group. Hence, we conclude this section with the observation that the raw data neither exclude nor confirm the existence of an effect of the 2019 smoking ban, and that we are required to construct a counterfactual that captures what would have happened in Sweden without the smoking ban to be able to perform a more thorough analysis. Since the smoking ban came into force nationally at the same time, there is no within-country variation in treatment assignment to be exploited in a statistical analysis. Hence, we turn to a cross-country analysis that uses untreated countries as a possible control group. Before pursuing such an analysis, we devote the next section to an investigation into the characteristics of Sweden and the control group countries to get a thorough understanding of what goes into the counterfactual outcomes to be estimated.

5.2 Characteristics of Sweden and the donor pool

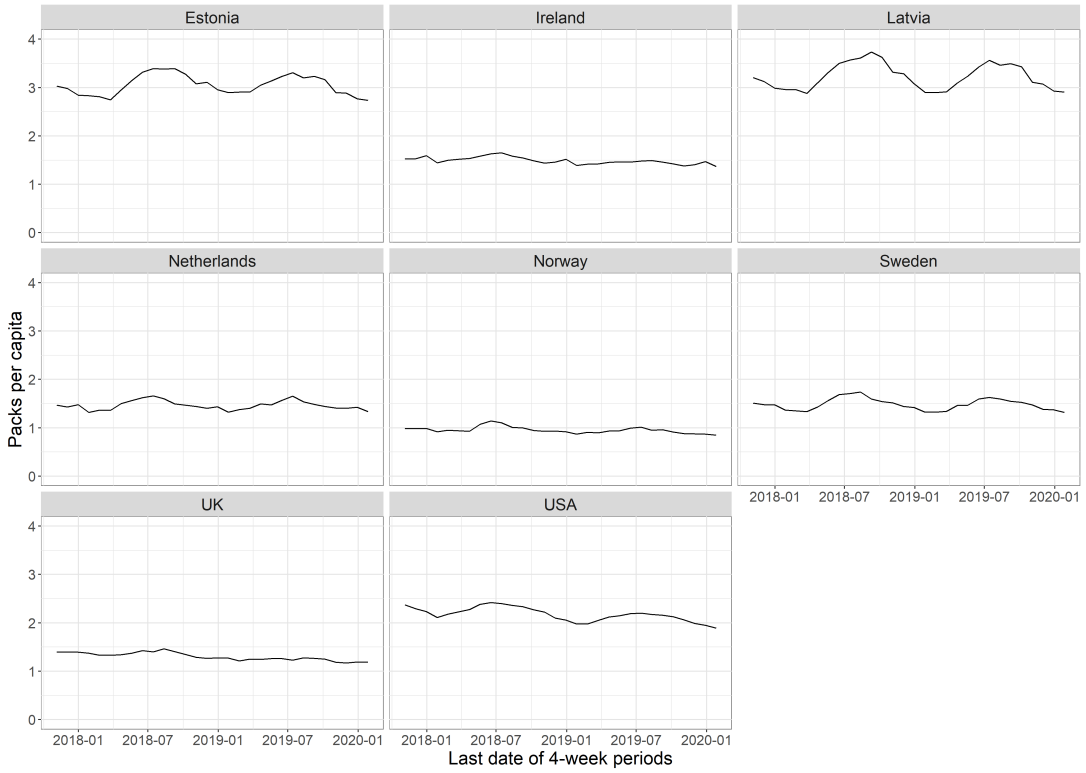
The countries in our donor pool are Estonia, Ireland, Latvia, the Netherlands, Norway, the United Kingdom and the United States (for details about the selection of these countries, see section 4.1). In this section, we descriptively investigate the characteristics of Sweden and the donor pool, to increase the transparency of our analysis with regards to what goes into the estimated counterfactual outcomes and to support a discussion of the external validity of our findings. Specifically, we investigate smoking demographics, the comprehensiveness of tobacco control policies, as well as climatic and weather differences that may be important determinants of the seasonality in cigarette sales that we observe.

5.2.1 Smoking consumption and demographics

We begin by taking a closer look at the outcome variable of interest—cigarette sales per capita—in all of the countries in our sample. In Figure 3, we plot the time series of monthly cigarette sales per capita for each country. It is worth repeating that the levels of the series are functions not only of the actual cigarette sales per capita but also of Nielsen's market coverage. Since Nielsen's market coverage varies in the range 68% to 91% across the countries (see Table 1), one should exercise caution when comparing the levels of the series. That being said, when ranking the countries by the level of their per capita cigarette sales, Sweden gets the third lowest rank using both monthly data and annual data based on tax records, with Norway and the United Kingdom having lower per capita consumption. This suggests that the graphs convey a fairly accurate picture about Sweden's *relative* level of cigarette sales. Furthermore, the graphs reveal an important characteristic of cigarette sales data, namely that seasonality is exhibited, which is a well-known phenomenon in the literature (see e.g. Chandra and Chaloupka 2003).

We now turn to annual cigarette sales data based on tax records in order to get a sense of the trends in cigarette sales over a longer time period. Figure 4 shows the annual number of cigarette packs sold per capita in each sample country between 2013

Figure 3
Monthly time series of cigarette packs sold per capita by country

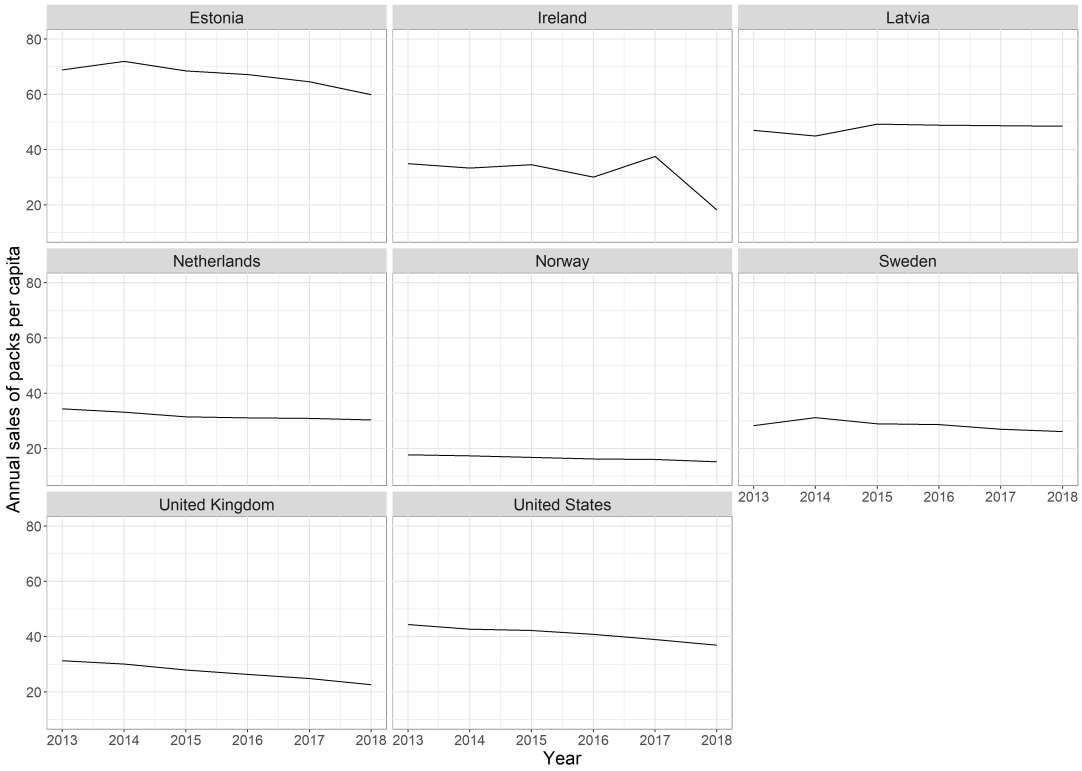


Source: Authors' rendering of data from Nielsen.

and 2018. As we can see, a majority of the countries exhibit rather similar downward trends. The exceptions are Latvia, whose trend appears to be mostly flat, and Ireland that saw rather dramatic spikes and slumps in 2017 and 2018, respectively. However, in the case of Ireland, this is an artifact of regulatory changes that lead retailers to stockpile cigarettes in 2017 that would otherwise have been reported as belonging to 2018 in the tax records (The Irish Department of Finance 2019). If we disregard these two years, Ireland's trend is rather similar to those of the other countries.

Next, we examine the characteristics of smokers in our sample countries. The data used in this section come mainly from a compilation of national surveys by WHO from 2018 complemented with data across age groups from the Europe-wide survey EHIS performed in 2014 (see section 4.3 for details). As seen in Figure 5, there is a fair amount of variation in the smoking prevalence in our sample, with the Baltic states of Latvia and Estonia in the upper end of the spectrum at around 30% of the population being current smokers, and the Nordic countries Sweden and Norway in the lower end of the spectrum at just below 15%, while the non-European United States, the northwestern European UK and Ireland, and the central European the Netherlands lie somewhere in between. This distribution is in line with the geographical differences that have typically been found in the literature (Joossens, Feliu, and Fernandez 2020). In contrast, and as previously mentioned, Sweden's per capita cigarette consumption is higher than Norway's

Figure 4
Annual time series of cigarette packs sold per capita by country



Source: Authors' rendering of data from European Commission (2020).

and the United Kingdom's and just below that of Ireland. The discrepancy between the two measures can partly be explained by the cross-border cigarette shopping of Norwegians—accounting for approximately 10% of the Swedish cigarette sales according to estimates (Zetterqvist and Ramstedt 2019)—being included in the Swedish sales statistics. However, after adjusting for the cross-border shopping, the Swedish per capita cigarette sales is still higher than that of Norway and the United Kingdom, although less so. Hence, it appears that while less Swedes report being smokers, those who do smoke consume cigarettes more intensively than in Norway and the United Kingdom.

Another insight from Figure 5 is that prevalence of occasional smokers relative to daily smokers varies substantially. Again, Sweden and Norway stand out, this time with a large share of the current smokers being occasional smokers. A possible explanation for this is that in Sweden and Norway, there are many tobacco users who have a smokeless substitute to cigarettes called snus as their primary source of nicotine but who may consume cigarettes in e.g. social contexts.[7] If we look instead at occasional smokers as a share of the adult population, the level is relatively more stable but still varies from a low of 1.5% in the UK to a high of about 5% in Estonia.

In Figure 6, we present the gender distribution of current smokers. Here, the Baltic

[7]This explanation has been promoted in previous literature, see for example Foulds et al. (2003) and Ramström and Foulds (2006).

Figure 5
Smoking by country and smoker type

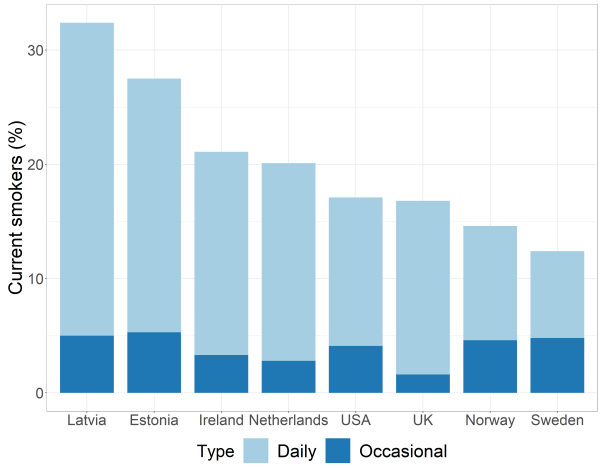
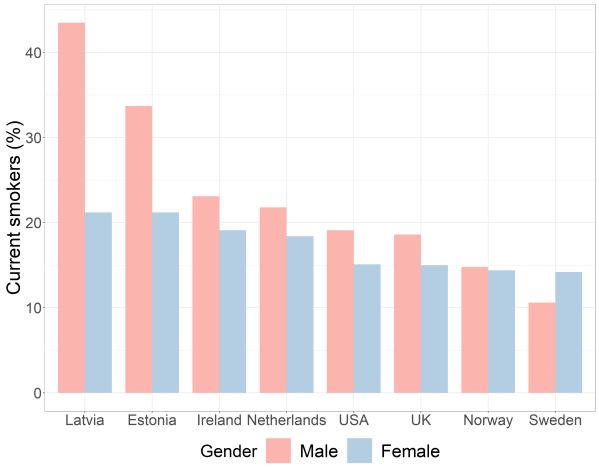


Figure 6
Smoking by country and gender



Note: The data come from surveys performed in year 2018. Source: Authors' rendering of data from WHO (2021).

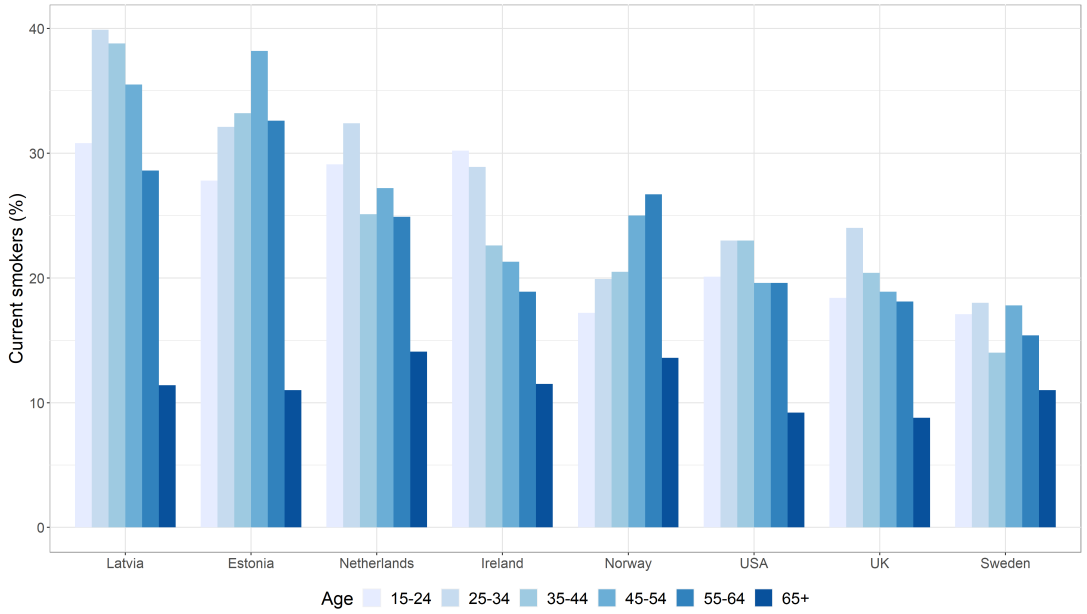
countries are outliers with substantially higher smoking prevalence among males than females, whereas the prevalence is similar across genders in the remaining countries. This could potentially point towards cultural differences with respect to smoking habits in the Baltics. In Figure 7, we plot the smoking prevalence by age group and country, and along this dimension it is more difficult to discern any distinctive patterns. While those aged over 65 years have the lowest smoking prevalence in all countries, the relative prevalence among younger age groups varies. In some countries the prevalence appears to increase with age, in some the opposite seems to be true and in some there is no clear relationship.

Based on the descriptive statistics presented in this section, it is evident that there is dispersion in the level of smoking across the sample countries, with Sweden being among the countries with the lowest level of smoking. However, this should not discourage further analysis using a cross-country methodology; it is not certain that differences in the *level* of smoking between Sweden and some of the control units render comparisons of the trend or seasonal pattern of cigarette sales invalid. Furthermore, the descriptive statistics reveal that Sweden stands out somewhat among the sample countries in the sense that a large share of the smokers are occasional smokers and that a smoke-free substitute to cigarettes is widely consumed. This might have implications for the external validity of our results, such that the smoking ban of 2019 might have had a different impact if it were implemented in a country with different characteristics, which we will come back to when discussing our results.

5.2.2 Tobacco control policies

Next, we briefly look into the regulatory status of tobacco control policies in the sample countries. As discussed in section 2.1, tobacco control policies are rather similar across Western countries due to the large degree of international collaboration within this policy

Figure 7
Smoking by country and age group



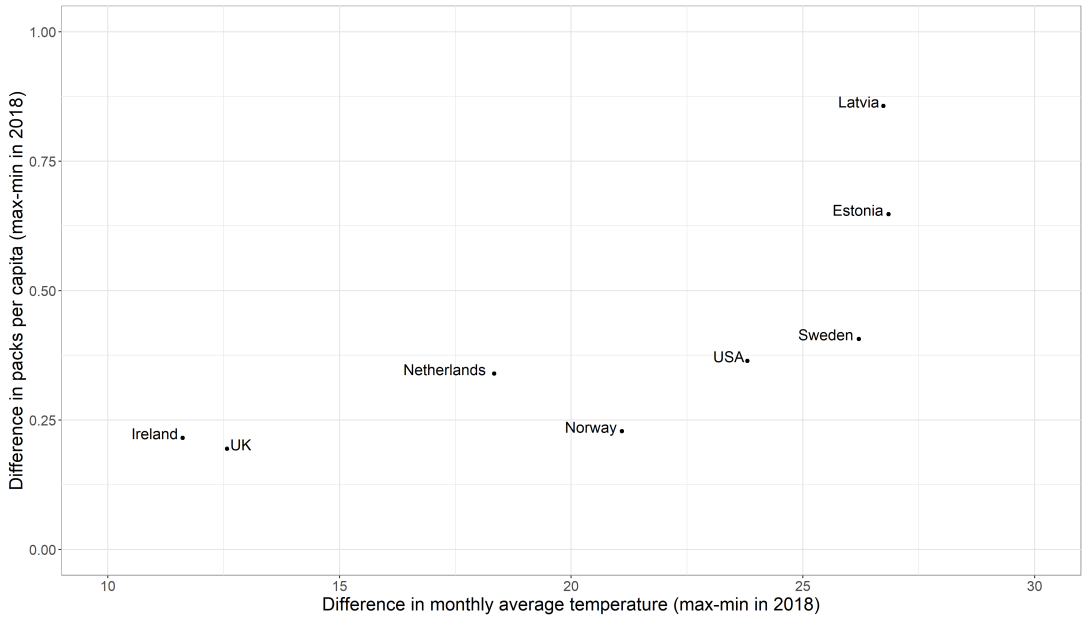
Note: USA data are only available for ages 18-24, 25-44, 45-64 and 65+. All data come from surveys performed in year 2014. Source: Authors' rendering of data from European Commission (2015b) and Hu et al. (2015).

area, and this is true for our sample countries at the time of the enforcement of the smoking ban of 2019 as well. Still, there was some heterogeneity in the comprehensiveness of the policies. While the tobacco control policies in Ireland, Norway and the United Kingdom were ranked as among the most comprehensive in Europe by the Association of European Cancer Leagues, the Netherlands and Sweden were in the middle group, and Latvia and Estonia were in the group with weakest tobacco legislation (Joossens, Feliu, and Fernandez 2020). Due to the US legislation being determined both at the federal and state/local level, but mainly at the latter, it is difficult to provide an overview of the tobacco control policies of the country as a whole. In Table A1 we present an overview of the major tobacco control policies that have been implemented in the countries during the last 20 years. Importantly, after excluding Austria, none of the countries have introduced any major tobacco control policies around the time of the enforcement of the Swedish smoking ban of 2019. Further, all sample countries have been continuously raising the cigarette tax rate in the years leading up to the treatment date, in an as good as linear manner. While there are some differences in these growth rates, with the average semiannual growth rate between 2017 and 2019 being 2.6% compared to Sweden's 1.4%, there are no indications of a trend shift in any of the countries around the time of the treatment date. Hence, we have no reason to believe that any of the control units experienced a policy-induced shock to cigarette sales around the time of the enforcement of the policy in study, which is an important criteria for identification.

5.2.3 Climate and weather

A final and possibly important characteristic to consider in relation to cigarette sales is climate and weather. In particular, as we saw in Figure 3, cigarette sales exhibit seasonality with peaks in the warmer months of the year and lows during winter. As discussed in Chandra and Chaloupka (2003), this seasonal pattern may be explained by a variety of factors coinciding with the seasons such as timing of tax changes (usually in December-January or June-July) or efforts to quit due to e.g. New Year's resolutions. However, a major candidate explanation is temperature and weather. While many still smoke indoors, predominantly within the bounds of the own home due to laws in most Western countries prohibiting indoor smoking in public spaces, it is to a large extent an outdoor activity. As such, the general pleasantness of smoking may well be affected by how agreeable the weather is, which would help explain the observed seasonal pattern in sales. Figure 8 plots the relationship between the difference in temperature between the warmest and coldest month and the difference in cigarette sales (packs per capita) between the most and least selling months, in 2018 over the countries in our sample. The positive relationship indicates that the magnitude of the temperature fluctuations over the seasons is important in explaining the degree of seasonality in cigarette sales, as countries with stronger temperature swings on average also seem to have a greater variability in cigarette sales over the year.

Figure 8
Relation between seasonality in temperature and cigarette sales



Source: Authors' rendering of data from Nielsen and Berkeley Earth (2021).

The reason why we look closer at this relation between weather and cigarette sales is that two major weather events took place in our study period, in the form of heat waves during the summers of 2018 and 2019. In the summer of 2018, the heat wave was

particularly severe in northwestern Europe, Scandinavia and parts of the United States (Vogel et al. 2019; Kueh and Lin 2020), while the heat wave in the summer of 2019 covered a larger part of Europe, but primarily centered around continental Europe in countries such as Germany, France and Austria (Zhao, Zhou, and Chen 2020). While temperatures above the seasonal mean over only a duration of a week or two is perhaps unlikely to affect cigarette consumption to a degree that would be clearly visible in our data, this begs the question of whether the countries in our samples experienced substantially different and persistent weather conditions over the summers of 2018 and 2019, in which case they may be deemed too different to include as control units for Sweden. Intuitively, the relation between cigarette sales and weather may imply that if Sweden saw an exceptionally warm summer in 2018 and a normal summer in 2019, we could get a good pre-treatment match between Sweden's actual series and the counterfactual series but still biased estimates of the treatment effect if the donor pool consisted of countries that instead had a normal summer in 2018 and an abnormally warm summer in 2019.

Figure 9 shows temperature series by country averaged over the same 4 week periods as our main data. Further, Figure 10 plots the mean temperature in June through August in 2018 and 2019 for the countries in our sample, where points below the 45 degree line are countries that on average had a colder summer in 2019 than in 2018, and vice versa. As we can see from these plots, all countries in our sample had higher temperatures in the summer of 2018 than in 2019. Moreover, the differences in mean temperature between the two summers among the countries are similar not only in direction, but also in magnitude. Hence, any heterogeneity in temperature shocks occurring over the two summers does not seem to be sufficiently large so as to be strongly noticeable when averaging over the three summer months. The Appendix Figure A1 shows the same differences but for the individual months of June through August, which shows a slightly higher heterogeneity in August. However, the direction of difference is in general uniform among all of the

Figure 9
Average monthly temperature by country

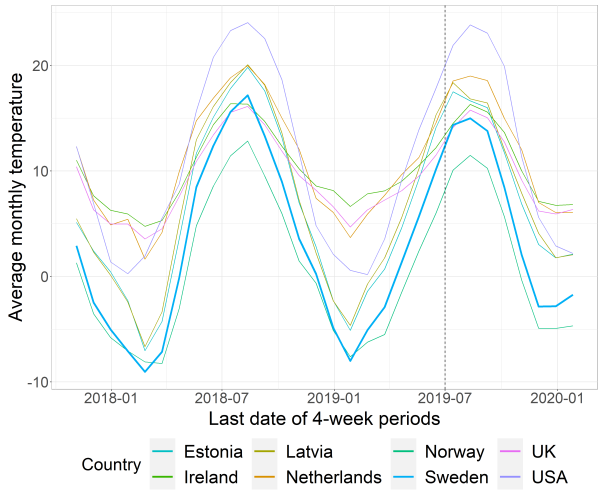
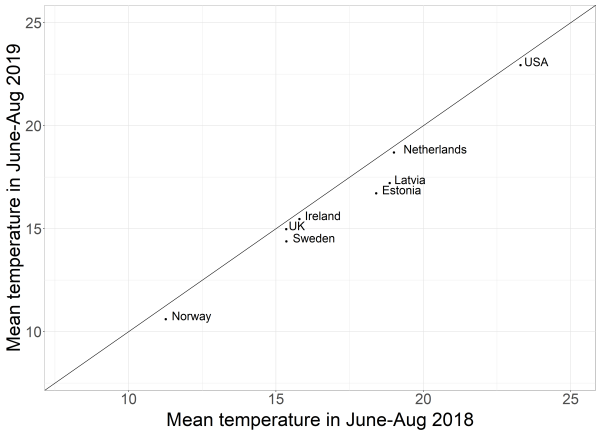


Figure 10
Summer temperatures in 2018 and 2019



Source: Authors' rendering of data from Berkeley Earth (2021).

countries. Given this, we believe that heterogeneity in temperature during these two heatwaves is not to be seen as a problem for our cross-country analysis. First, even though the strength of seasonal fluctuations in temperature correlates positively with the variability of cigarette sales over the year, it is not certain that temperature shocks *above and beyond* the seasonal mean of a given country affects smoking to any noteworthy degree. To be best of our knowledge, this subtle but important difference in the impact of weather on smoking has not yet been studied in the literature. Second, even if such an effect would exist, the heterogeneity in temperature over the two heat waves seems to be small, as previously discussed. In the absence of heterogeneity, the heat waves would amount to common shocks affecting all countries equally, in which case they do not present a problem unless the effect on smoking itself is heterogeneous. However, this last point is difficult to address given the aforementioned lack of studies on this topic.

In this section, we have provided an overview of the characteristics of the countries in our sample. While differences exist, particularly with regards to the level of smoking across countries, we believe that they do not invalidate a cross-country analysis of the 2019 smoking ban in Sweden, as previously discussed. In the next section, we present the first method of analysis that we will employ in our analysis—the SC method.

5.3 The Synthetic Control Method

5.3.1 Conceptual framework

In order to facilitate the exposition of the methods used in this paper, and in particular the differences between the SC and MC estimators, we provide a brief conceptual framework on how they function before going into formal theory. In this section, we focus on the SC method.

Matrix $\mathbf{M}$ is a hypothetical matrix in which each row represents outcomes of a single unit, and each column denotes a particular point in time.[8] Hence, an entry is the value observed for a given unit at a given time. In this matrix, observed and untreated data points are denoted by the symbol ✓, while unobserved counterfactual data points for the treated unit in the last row are denoted by question marks. The fundamental issue at hand is to estimate the question marks, since they capture what would have happened in the absence of treatment.

$$\mathbf{M} = \begin{bmatrix} ✓ & ✓ & ✓ & ✓ & \cdots & ✓ \\ ✓ & ✓ & ✓ & ✓ & \cdots & ✓ \\ ✓ & ✓ & ✓ & ✓ & \cdots & ✓ \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ ✓ & ✓ & ✓ & ✓ & \cdots & ✓ \\ ✓ & ✓ & ✓ & ? & \cdots & ? \end{bmatrix}$$

[8]This exposition draws heavily on Athey et al. (2021).

The SC method takes the observed values for the control units (all ✓'s in the pre-treatment period for all rows except the last) as given, and tries to construct a new last row that resembles the observed values for the treated unit by means of a weighted average of the control unit rows. This constructed row is known as the synthetic control unit, and is the estimate of the counterfactual outcomes for the treated unit in the absence of treatment. Hence, conceptually, the SC method searches for patterns across units that are stable over time in order to construct a new series that matches the pre-treatment series of the treated unit.

5.3.2 Theory

In this section, we describe the SC method as formulated in Abadie (2020). The setting is such that we observe data for $J+1$ units ($j = 1, 2, ..., J+1$) where $j = 1$ is the treated unit and we refer to $j = 2, 3, ..., J+1$ as our "donor pool". Further, we observe these units over T ($t = 1, 2, ..., T$) time periods, whereof the first T_0 observations are measured before the intervention. We observe our outcome variable of interest, Y_{jt}, for each combination of unit j and time t, and a set of k predictors of the outcome variable, $X_{1j}, X_{2j}, ..., X_{kj}$, for each unit j. In terms of notation, the $k \times 1$ vectors containing the set of predictors for units $j = 1, 2, ..., J+1$ are denoted $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_{J+1}$, and the $k \times J$ matrix in which the predictor vectors for the donor pool units are stacked is denoted $\mathbf{X}_0 = [\mathbf{X}_2\ \mathbf{X}_3 \cdots \mathbf{X}_{J+1}]$.

Next, we describe the statistical problem that the SC method is intended to solve. Following Abadie (2020) and using the terminology of Rubin's Model for Causal Inference, Y_{jt}^N is the potential response for unit j at time t in the absence of treatment and Y_{jt}^I is the potential response for a treated unit j in a post-treatment period. We can now express the treatment effect, allowing the effect to vary across post-treatment periods $t > T_0$, as

$$\tau_{1t} = Y_{1t}^I - Y_{1t}^N. \tag{1}$$

At this point, it should be noted that for all post-intervention periods, we only observe $Y_{1t} = Y_{1t}^I$ for the treated unit. In other words, the potential outcome of the treated unit in the absence of treatment is unknown to us. This problem is commonly known as the fundamental problem of causal inference (Rubin 1974; Holland 1986). Estimating the treatment effect is because of this problem equivalent to estimating the potential outcome in the absence of treatment for the treated unit, Y_{1t}^N, also known as the counterfactual outcome. As described in the previous section, the SC method is a method for retrieving an estimate of the counterfactual outcome. It is based on the idea that in many cases, a combination of the donor pool units may be a better approximation of the treated unit than any of the individual control units. In practice, a synthetic control for the treated unit is constructed as a weighted average of units from the donor pool, and can be characterized by the $J \times 1$ vector of weights $\mathbf{W} = (w_2, w_3, ..., w_{J+1})$. The synthetic

control estimators of Y_{1t}^N and τ_{1t} are

$$\widehat{Y}_{1t}^N = \sum_{j=2}^{J+1} w_j Y_{jt}, \tag{2}$$

and

$$\widehat{\tau_{1t}} = Y_{1t} - \widehat{Y}_{1t}^N. \tag{3}$$

The procedure of choosing the elements of $\mathbf{W}$ proposed by Abadie and Gardeazabal (2003) follows two steps. In the first, the set of weights in $\mathbf{W}$ is chosen such that the synthetic control unit resembles the treated unit in terms of its predictor values, $\mathbf{X}_1$. Here, a set of non-negative constants $\mathbf{V} = (v_1, v_2, ..., v_k)$ expressing the relative importance of each of the k predictors is taken as given. In other words, given $\mathbf{V}$, the synthetic control $\mathbf{W}^* = (w_2^*, w_3^*, ..., w_{J+1}^*)$ is chosen to minimize

$$||\mathbf{X}_1 - \mathbf{X}_0\mathbf{W}|| = \left(\sum_{h=1}^{k} v_h(X_{h1} - w_2 X_{h2} - \cdots - w_{J+1}X_{hJ+1})^2\right)^{1/2}. \tag{4}$$

The most common version of the SC method also constrains the weights $(w_2, w_3, ..., w_{J+1})$ to be non-negative and sum to one. By doing so, one avoids extrapolation and reduces the risk of overfitting on complicated, spurious cross-country relationships. It should be noted, however, that a degree of extrapolation will still be present in our analysis since we transform the data to percentage deviations from the pre-treatment mean, which centers the data at zero.

From (4), we find the optimal country weights $\mathbf{W}^*$ conditional on a vector of predictor weights $\mathbf{V}$. However, in order to find the optimal $\mathbf{V}^*$, Abadie (2020) proposes a cross-validation procedure. In particular, in the second step of the procedure the following sum is minimized with respect to $\mathbf{V}$, the predictor weights, over a *subset* of pre-intervention periods denoted $\mathcal{T}_0$:

$$\sum_{t \in \mathcal{T}_0} \left(Y_{1t} - w_2(\mathbf{V})Y_{2t} - ... - w_{J+1}(\mathbf{V})Y_{J+1t}\right)^2, \tag{5}$$

where the notation $w_j(\mathbf{V})$ highlights the fact that the optimal weight w_j depends on the predictor weights $\mathbf{V}$, and so on for all elements in $\mathbf{W}$. Once the optimal vector $\mathbf{V}^*$ is found through cross-validation, it is straightforward to obtain the optimal country weights $\mathbf{W}^*$ from $\mathbf{W}(\mathbf{V}^*)$. With this result, the synthetic control unit is constructed by the formula given in (2).

5.3.3 Analysis and results

In this section, we present results from the SC method along with a number of placebo tests, after motivating our choice of outcome variable transformation and predictors. Throughout the section, the dimension of the data is N=8, where one of the countries is the treatment unit Sweden, and T=30, where the last eight periods are treated. This leaves us with 22 pre-treatment periods. In all specifications, the last eight pre-treatment periods are held out as a validation set in the cross-validation procedure.

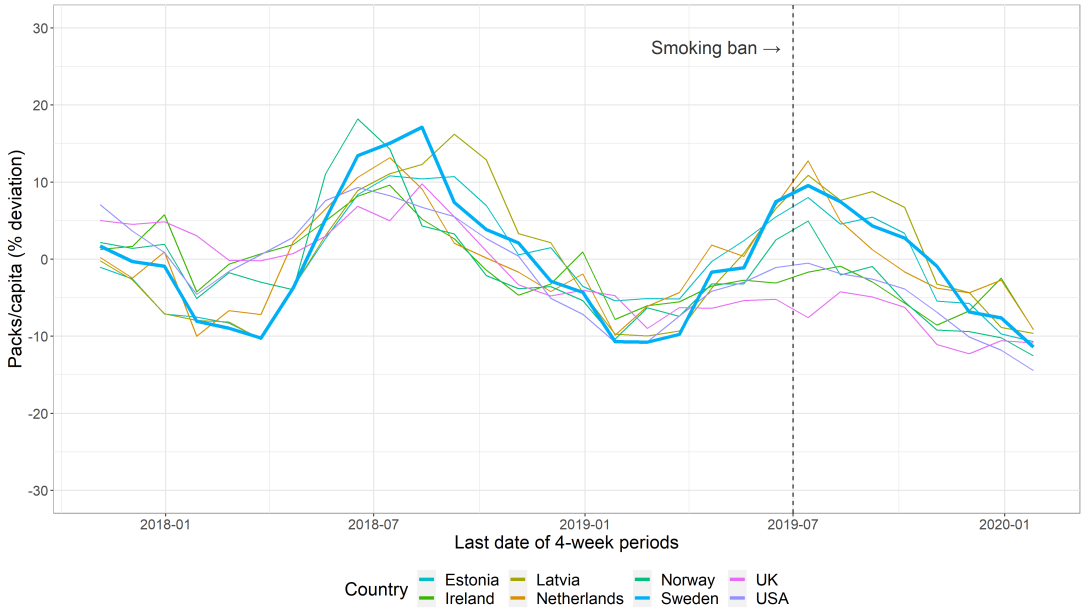
The outcome variable used in this and the rest of the main analysis is cigarette pack sales per capita, transformed into percentage deviations from the pre-treatment mean (shown by country in Figure 11). While it would have been technically possible to perform the analysis with the outcome variable expressed in levels since Sweden lies within the convex hull of the donor pool units, we consider it is undesirable to do so. There are several reasons for this.[9] To begin with, we deem it necessary to center the data at zero since the level differences between countries will not only reflect true differences in cigarette sales but also the varying degree of Nielsen's market coverage in different countries, and we cannot reliably estimate these market coverage rates.[10] In addition, we argue that since we are studying short-term effects, matching on structural determinants of smoking levels is less important than it would have been if we studied effects over many years, where e.g. floor effects may dampen the long-term trend in countries starting at an already low level of smoking. A possible remedy in situations when matching on level is not appropriate, as suggested in Abadie (2020), is to demean the series with respect to their pre-treatment means by country. However, the cyclical patterns observed in our data strongly suggest that cigarette consumption varies over the seasons relative to the level of baseline consumption, and not by a fixed amount across all countries. Hence, simply demeaning the series will introduce larger fluctuations in absolute terms for countries with a high level of smoking, and vice versa for low level countries, although the percentage changes may in fact be highly comparable. Therefore, we transform the series into percentage deviations from the pre-treatment mean. It is important to note that such a transformation changes our identifying assumption. The counterfactual that we estimate does not need to be able to replicate the level of Swedish cigarette sales per capita, but rather the degree of its seasonal swings relative to some baseline consumption and any time trends present over the study period. As such, we assume that the level of synthetic Sweden is unimportant for its ability to predict Sweden's outcomes in the absence of treatment.

Our choice of predictors used in the main specification is motivated next. As per the discussion above, we want the estimated synthetic Sweden to match actual Sweden in

[9] For a more thorough discussion on this, see Appendix B.

[10] Imagine, for example, that we apply the SC method on cigarette sales in levels and hence include predictors that are informative for this measure. However, since the market coverage of Nielsen varies, the measured level of cigarette sales also varies in ways not captured by the predictors. Hence, the structural link that may exist between the predictors and the true outcome will not necessarily exist in the data we use, such that the predictors become rather uninformative. Indeed, this is confirmed in Section 6 where we perform this analysis as a robustness check.

Figure 11
Monthly time series of cigarette packs sold per capita after transformation by country



Source: Authors' rendering of data from Nielsen.

terms of its cyclical behavior of and trend in sales. Therefore, a set of predictors that can predict the trend and seasonality of cigarette sales should be included. This step is important since the period of our outcome data is too short for us to be able to formally estimate any long-term trend and seasonal components of the series. As predictors of the trend in cigarette sales, we include the mean growth in annual sales of cigarette packs per capita between 2013 and 2018,[11] and the growth in tax rates on cigarettes between January 1, 2017, and January 1, 2019. The former is chosen to allow us to better capture the trend in the data, since it is arguably difficult to do so with the 22 pre-treatment months in our data set, while the latter captures changes in the price of cigarettes that are exogenous with respect to cigarette demand. For the seasonal component, we include two predictors that are possibly important for determining the degree to which cigarette sales fluctuate over the year. The first one is the percentage of occasional smokers by country, measured in 2018, as this group is arguably more likely to change their consumption based on external factors that change with the seasons than daily smokers, due to their smoking behavior being contextually dependent. Second, we include a measure of the average temperature difference between the warmest and coldest month over the years 2006–2016 in each country. This is meant to capture the average strength of the seasonal fluctuations in weather and climate, which our investigation in section 5.2.3 suggests is an important determinant of the seasonality in cigarette sales. Finally, as is standard practise in the literature (see e.g. Abadie, Diamond, and Hainmueller 2015), we include a few transformations of the outcome variable in the pre-treatment period: the means over

[11] Due to the previously discussed regulation in Ireland distorting the tax data in 2017 and 2018, we exclude these years for Ireland only in the calculation of the mean sales growth.

Jan-Aug 2018, Dec–Feb 2018–19 and Apr–Jun 2019 respectively. One could also have considered including income and cigarette price as predictors, since such variables are known to be important determinants of demand of consumer goods (Abadie, Diamond, and Hainmueller 2010). However, while including some measure of the *level* of income and price would be important if our outcome variable was also expressed in levels, we expect their predictive power to be highly limited in the case of our transformed outcome variable. In particular, even if the level of income and price could for example be argued to affect the composition of smoking types, we can account for such differences directly by including the share of occasional smokers as a predictor. Hence, we do not use income and price levels as predictors. An alternative would be to include the growth in income and prices over the pre-treatment period. However, it appears improbable that GDP growth would affect cigarette demand over the short time horizon of our analysis. Accordingly, when growth in GDP and price (instead of tax rates, one of the main components of cigarette prices) were included as predictors, GDP received a very small predictor weight and the overall impact on the results were negligible since the changes in prices are well captured by changes in tax rates.

We now move on to presenting the results. Figure 12 plots the actual series (Y_1^I) and the counterfactual estimated with the SC method ($\widehat{Y}_1^N$), and Figure 13 shows the estimated treatment effect ($Y_1^I - \widehat{Y}_1^N$). As we can see here, the pre-treatment fit is quite good overall but misses the peak in 2018 by some margin. Comparing the fit to the patterns in Figure 11 we see that Sweden has among the highest peaks—with the peak in 2018 being an outlier relative to the donor pool series—and lowest minimums. This explains synthetic Sweden's inability to fully capture Sweden's seasonal pattern since the version of the SC method used here does not allow for extrapolation in the sense that the country weights are constrained to non-negative and sum to unity. Looking at the post-treatment periods, Figure 12 shows no sign of any treatment effect as the counterfactual closely tracks the actual cigarette sales per capita that is observed in Sweden. However, given the poor fit during the 2018 summer, it is questionable if synthetic Sweden can be seen as a good counterfactual for the summer of 2019, which we will come back to when discussing the interpretation of our results.

As can be seen in Table 2, the counterfactual, or synthetic, Sweden is a weighted combination of Latvia, the Netherlands, Norway and the United States with zero weight given to the other donor pool countries. Latvia is especially important with its weight of 0.503. Further, Table 3 shows that the cross-validation procedure gives a fairly large weight (0.428) to the outcome mean in the periods right before the treatment, which is a common feature of the SC method. Interestingly, the average seasonal temperature difference is given a weight as high as 0.227, which lends further support to the suspicion that weather fluctuations are important in explaining the seasonality in cigarette sales as previously discussed. Mean growth in tax rates 2017–2019 also appears to be an important predictor for the cigarette sales series with a weight of 0.242. Table 4 shows the predictor balance between the treated and synthetic unit, as well as the means in the entire sample. In all of the predictors, the synthetic unit captures the characteristics of

Figure 12
Actual and Synthetic Sweden (SC)

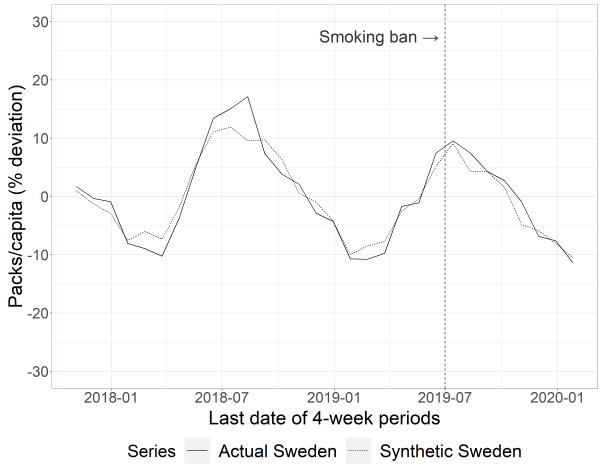


Figure 13
Treatment effect (SC)

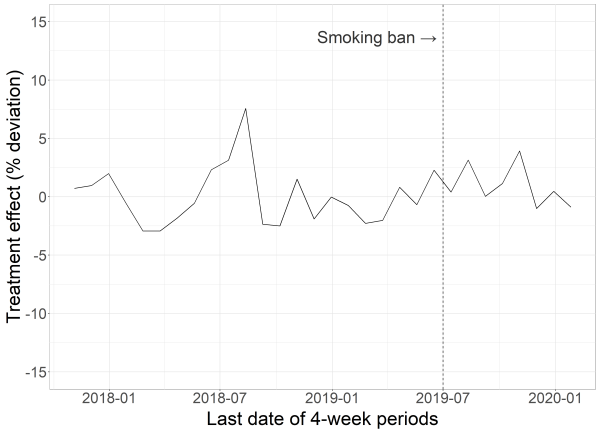


Table 2
Country weights

	W
Latvia	0.503
Netherlands	0.214
Norway	0.195
USA	0.088
Estonia	0
Ireland	0
UK	0

Table 3
Predictor weights

	V
Seasonal temperature difference, 2006–2016 (°C)	0.227
Mean growth in cigarette sales, 2013–2018 (%)	0.009
Prevalence of occasional smoking, 2018 (%)	0
Mean growth in tax rate, 2017–2019 (%)	0.242
Outcome mean, Jun–Aug 2018	0.009
Outcome mean, Dec–Feb 2018–19	0.085
Outcome mean, Apr–Jun 2019	0.428

Table 4
Predictor balance

	Treated	Synthetic	Sample mean
Seasonal temperature difference, 2006–2016 (°C)	23.447	21.179	18.458
Mean growth in cigarette sales, 2013–2018 (%)	-1.441	-1.059	-3.142
Prevalence of occasional smoking, 2018 (%)	4.800	4.372	3.814
Mean growth in tax rate, 2017–2019 (%)	1.427	1.851	2.575
Outcome mean, Jun–Aug 2018	15.210	10.872	9.551
Outcome mean, Dec–Feb 2018–19	-5.954	-5.065	-4.822
Outcome mean, Apr–Jun 2019	1.565	0.758	-0.856

Source: Data on seasonal temperature difference 2006-2016 is collected from World Bank (2021a). Mean growth in cigarette sales 2013-2018 is calculated using data from European Commission (2020). Data on prevalence of occasional smoking 2018 is collected from WHO (2021). Mean growth in tax rate 2017-2019 is calculated using data from European Commission (2021). Outcome means Jun-Aug 2018, Dec-Feb 2018-19 and Apr-Jun 2019 are based on data from Nielsen.

Sweden better than the sample mean would, and in many cases significantly better. Still, the matching is not perfect, although we consider it fairly good given the rather small size of our donor pool. As we expected, and as we saw in the plot of actual and synthetic Sweden, there is a substantial deviation in the outcome mean in Jan-Aug 2018 which is the period where Sweden deviates the most from the donor pool countries. Figure A2 in the Appendix shows the sensitivity of the estimated synthetic unit to the exclusion of each predictor separately. In general, the impact of excluding a certain predictor is very small.

The results above suggest that, in terms of the chosen structural characteristics included in the specification, the synthetic control unit does much better in capturing actual Sweden than an unweighted average of the donor pool countries would have. However, we should still be careful in drawing strong conclusions about the obtained null result since the deviation of synthetic Sweden from actual Sweden in the summer of 2018 is non-negligible. If a similar deviation occurred in the summer of 2019, we could be missing an actual treatment effect in the expected direction.

In what follows, we will carry out the robustness checks that are the standard in the SC literature. In particular, it is important to investigate what a typical prediction error looks like. If the variance of the predictions is large, outweighing any plausible treatment effect, we would run the risk of being underpowered. The standard approach in the literature is to run two kinds of placebo tests to address this potential issue, allocating treatment assignment to the other countries in the donor pool (country placebos) and assigning treatment status at an earlier point in time for the treated unit (in-time placebos). We will show the results from these tests below, as well as a combination of both of them.

For the country placebos, we follow the convention of excluding countries that have a sufficiently bad pre-treatment fit as measured by the mean squared prediction error (MSPE) of the estimated counterfactual unit over the entire pre-treatment period. There are no clear guidelines on how to choose this threshold, and in particular, excluding too many countries when the donor pool is already fairly small will make it more difficult to draw any conclusions about the typical post-treatment prediction error. Abadie, Diamond, and Hainmueller (2010) uses a threshold of two times the MSPE of the treated unit in their final specification, but we choose to be slightly more lenient due to the size of our donor pool and set the threshold to three. Under this restriction, no countries are excluded from the placebo testing. The in-time placebo test is more straightforward. In the main analysis, the number of treatment periods to be predicted is eight. Therefore, we simply reassign the treatment date to eight periods before the actual treatment date, which corresponds to December 1, 2018.[12] This test gives an indication of how well we can reproduce the Swedish cigarette packs per capita over an untreated prediction period that is equally long as and takes place just before the actual treatment period.

The country placebo tests are shown in Figure 14, and the in-time placebo is shown in Figure 15. The pre-treatment fit that we observed for Sweden seem to be rather

[12]Since a period in our data refers to a 4-week period, eight such periods correspond roughly to seven calendar months.

Figure 14
Country placebo tests (SC)

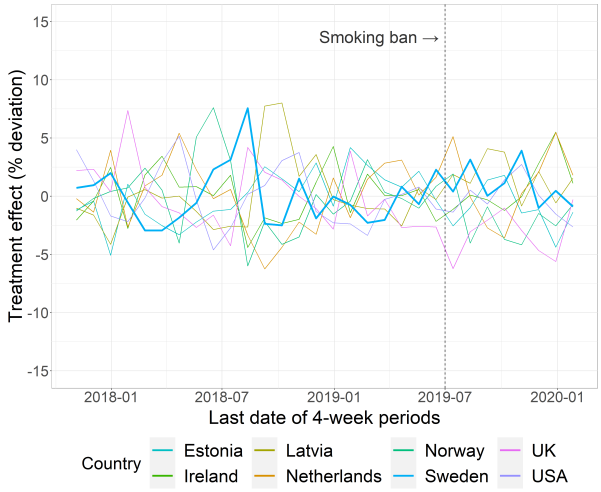
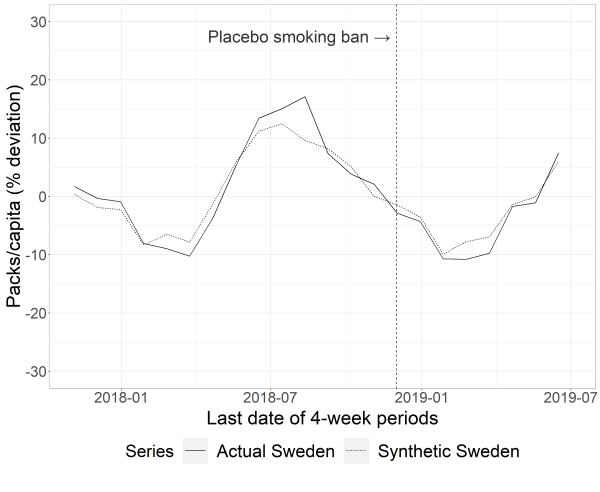


Figure 15
In-time placebo test (SC)



typical for the countries in our sample, as most of them fluctuate within the 5 percentage point interval. Further, the prediction errors in the post-treatment periods are also rather similar, following the same bandwidth between -5 to 5 percentage points. Since the number of donor pool countries is rather small, we should not interpret this as a confidence interval per se. However, it gives an indication of to what extent we may plausibly regard a deviation of a certain size as a treatment effect rather than just noise. In the in-time placebo test, synthetic Sweden manages to track the realized outcomes for Sweden very well, indicating that the model does quite well in predicting an accurate counterfactual for the treated unit. In order to make more use of the data we have, we can also combine the in-time and country placebos by allocating a placebo treatment at the earlier date *and* to a different country that was, in fact, not treated. This is shown in Figure 16. The prediction errors all seem to fall, with some margin, into the same bandwidth of -5 to 5 percentage points as before.

As previously noted, the country weights in our model are sparse, as is typical for the SC method. While this has the advantage that the source of the estimated counterfactual becomes transparent and easy to interpret, it poses a question of how sensitive the estimates are to the inclusion or exclusion of certain control units. Given the limited size of our donor pool, it is especially important to pay attention to this potential issue. To test whether this is a problem in our case, we perform a fourth placebo test that is commonly called a leave-one-out test. As suggested by the name, we perform the test by re-estimating the model $N-1$ times, each time leaving one of the donor pool countries out of the analysis. The results from this test can be found in Figure 17, where we plot the gap between the actual and estimated outcomes of Sweden (in the same way as in Figure 13) for each of the re-estimated models. While the general pattern is rather consistent across specifications, the pre-treatment fit becomes significantly worse when Latvia is excluded. Given the weight that Latvia is given in the main specification, this is not

Figure 16
In-time country placebo tests (SC)

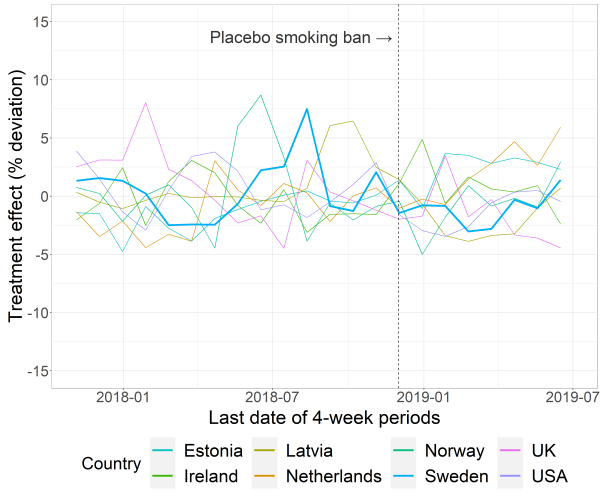
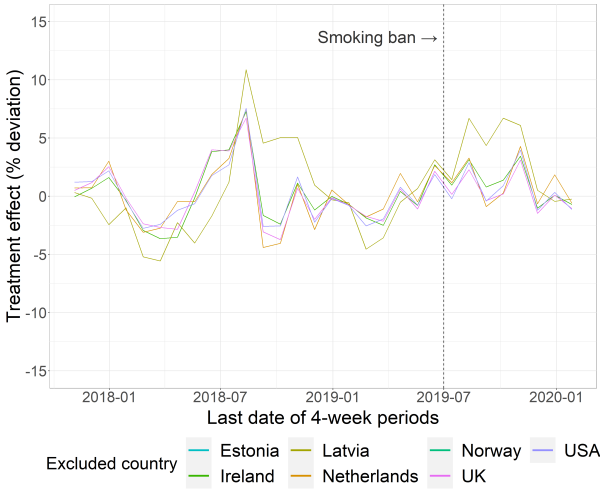


Figure 17
Leave-one-out analysis (SC)



too surprising. Further, when excluding Latvia, the treatment effect becomes positive at around the 5% mark for many of the post-treatment periods. Since we have little reason to believe that the true treatment effect is positive, this is likely due to the fact that without Latvia, Sweden is even more of an outlier in the data particularly during the summer months. This is further supported by the fact that the post-treatment periods in which excluding Latvia has a large effect corresponds to the months of the year (in 2019) for which the model also get a worse pre-treatment fit (in 2018). This sensitivity of the model to the inclusion of Latvia is, again, likely a problem that arises because we do not allow for extrapolation.

The general conclusion from our SC model is that while we fail to detect any effect of the smoking ban of 2019, the questionable fit of synthetic Sweden during the pre-treatment summer of 2018 warrants some caution in interpreting the null result as conclusive evidence of the policy not having any effect. We also conclude that the the poor pre-treatment fit, but also the sensitivity of the model to the inclusion of Latvia, is likely driven by our choice not to allow for any extrapolation. In order to further investigate this, we now turn to an alternative version of the SC method that relaxes the constraints on the country weights.

5.3.4 Relaxing the country weight constraints

As noted in the previous section, the synthetic control unit has some difficulties with capturing the relatively large fluctuations between winter and summer present in the outcome series of Sweden. This was not unanticipated since the descriptive plot of Figure 11 clearly shows that Sweden is an outlier in certain periods and under the default SC constraints on the country weights, restricting them to sum to unity, it is impossible for the method to provide a good fit in these periods. This constraint is a way of safeguarding

against extrapolation outside of the convex hull of the donor pool outcomes. However, as discussed in e.g. Doudchenko and Imbens (2016), allowing for extrapolation is not necessarily a violation of any identifying assumptions per se but should be determined on a case-by-case basis; allowing for an intercept in a DID analysis is, for example, often regarded as admissible despite the extrapolation it introduces. This is also brought up in the practical summary of the SC methodology by Abadie (2020). Seeing as Sweden is a slight but not in any way extreme outlier in a few periods, allowing for a certain degree of extrapolation may be warranted in this application. In particular, this would allow the synthetic control unit to give a larger weight to countries that have similar patterns in terms of timing of highs and lows as Sweden, but whose seasonal fluctuations are not as strong.

In this section, we consider an alternative set of constraints where the country weights are not restricted to sum to unity, while still preserving the constraint that individual country weights must be non-negative. Having a sparse set of country weights is an attractive feature of the SC method as is simplifies the interpretation of the synthetic control unit. Therefore, we consider the non-negativity constraint to be important in our setting. Moreover, since Sweden is not an extreme outlier in our data, there is no justification for allowing extreme degrees of extrapolation that would increase the risk of overfitting on spurious patterns in the data. As such, it would be an undesirable feature of this relaxed model if the country weights were to sum to a value very far from unity. It turns out that the solution chosen by the algorithm under these relaxed constraints is a country weight vector that sums to 1.457, which we deem acceptable. An alternative way of ensuring that the optimal country weight vector is not too complex would be to use regularization techniques as proposed in Doudchenko and Imbens (2016).[13] However, we see little reason to go into such territory. To begin with, regularization is particularly important when the size of the donor pool is large, in which case it safeguards against a non-sparse weight vector. With our rather small donor pool, this is not a problem. Moreover, regularization may still be important if we were to greatly relax the constraints, for example by allowing for negative weights, but as we have already discussed, such a relaxation seems superfluous in our case.

Relaxing the constraint that the country weights must sum to unity has important implications for how the SC method should be appropriately applied, in particular with regards to predictors. Consider, for example, if one were to include a predictor expressed in percentage terms, such as the share of young people in the population, while allowing the sum of the country weights to exceed unity. In practise, this may lead to situations in which this predictor value is above 100% for the synthetic unit, which clearly does not make sense. The recommendation when allowing for extrapolation as well as covariates is to regress the outcome variable on the pre-treatment predictors by simple OLS, and

[13]Regularization is the practise of including a term in the objective function of the optimization problem that penalizes the algorithm for choosing an excessively complicated solution. Common examples include attaching some positive constant to the l_1 or l_2 norms (i.e. LASSO or Ridge regression, respectively) to the vector of estimated coefficients or, in our case, the country weights.

then estimate the SC method on the obtained residuals using the entire vector of pre-treatment outcomes as predictors (Doudchenko and Imbens 2016; Abadie 2020).[14] We follow this recommendation and use the same predictors as in the default SC method in the OLS regression in order to obtain the residuals.

Figure 18 and 19 show the outcome series of Sweden and the synthetic control unit, as well as the treatment gap, when the model is estimated without the constraint that country weights must sum to unity. Allowing for extrapolation leads to a strong improvement of the pre-treatment fit, in the summer peak of 2018 but also in the lows, yet the results still indicate a null effect as synthetic Sweden tracks the series of actual Sweden fairly well in the post-treatment periods. If anything, the counterfactual series lies slightly below the actual series, which would indicate a treatment effect in the opposite direction of what we would expect. Table 5 shows the country weights. As previously mentioned, the optimal solution under the relaxed constraints involves an upward extrapolation, and the sum of the weights is equal to 1.457. We note that the countries that received a large weight in the previous SC specification, Latvia and Netherlands, are still important but now with the addition of UK and Estonia. Further, the desirable property of the weight vector being rather sparse is maintained.

Further, we run the same in-time, country and in-time country placebo tests, as well as the leave-one-out analysis, as before. These results are shown in the Appendix Figures A3-A6. Overall, the tests yield highly similar results as the main analysis, with most prediction errors in the placebo tests lying within the bounds of $\pm 5\%$. However, the leave-one-out analysis shows that the synthetic control unit for Sweden is now much less reliant on the country given the largest weight in the default version—Latvia. This is reasonable since the permitted extrapolation expands the set of feasible country weight vectors over which the algorithm searches for its solution. Moreover, it indicates that the large weight assigned to Latvia in the original specification was mainly motivated by Latvia's large seasonal fluctuations being suitable for making synthetic Sweden's fluctuations comparable to actual Sweden's in terms of magnitude.

We conclude this section by noting that the pre-treatment fit of the synthetic control unit is improved substantially when the constraint on the sum of the country weights is relaxed. In our case, this leads to a degree of extrapolation, but it does not seem excessive. The takeaway from the results in this section are virtually the same as before; there are no indications of an effect of the 2019 smoking ban. In the two versions of the SC estimator that we have presented, we have varied the constraints under which the SC estimator operates, while the fundamental workings of the model have remained unaltered. In the next section, we turn to a different approach to estimating the counterfactual outcomes of Sweden based on matrix completion techniques, which has the advantage of making a greater use of the full correlation structure of the data that we have.

[14]Note that including the entire vector of pre-treatment outcomes, instead of transformations of it such as the mean over certain periods, was not an alternative in the default SC estimator when other predictors were also included. As shown in Kaul et al. (2015), doing so renders all other predictors other than the pre-treatment outcome vector irrelevant.

Figure 18
Actual and Synthetic Sweden (relaxed SC)

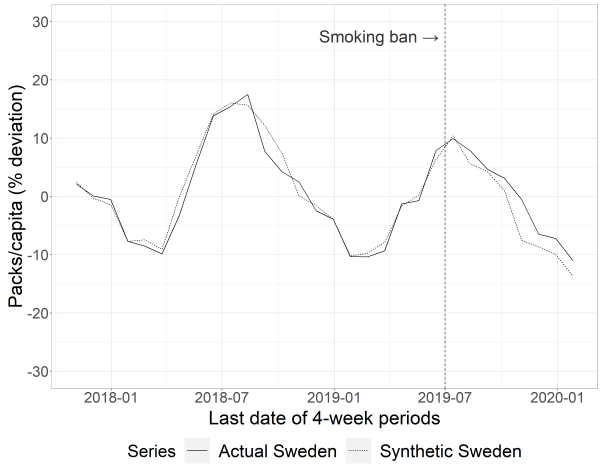


Figure 19
Treatment effect (relaxed SC)

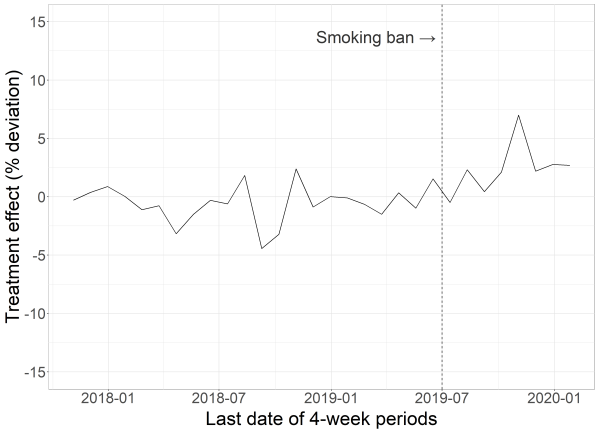


Table 5
Country weights (relaxed SC)

	W
Netherlands	0.469
UK	0.334
Estonia	0.330
Latvia	0.296
Norway	0.028
Ireland	0
USA	0
Sum	1.457

5.4 Matrix Completion with Nuclear Norm Minimization

5.4.1 Conceptual framework

We begin by outlining the conceptual framework of the MC estimator by using the same example as we did for the SC estimator in section 5.3.1. Again, we exemplify with a hypothetical matrix $\mathbf{M}$ where the rows and columns denote unit and time respectively. Observed untreated data points are, again, denoted by the symbol ✓ while counterfactual outcomes of interest are denoted by question marks.

$$\mathbf{M} = \begin{bmatrix} ✓ & ✓ & ✓ & ✓ & \cdots & ✓ \\ ✓ & ✓ & ✓ & ✓ & \cdots & ✓ \\ ✓ & ✓ & ✓ & ✓ & \cdots & ✓ \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ ✓ & ✓ & ✓ & ✓ & \cdots & ✓ \\ ✓ & ✓ & ✓ & ? & \cdots & ? \end{bmatrix}$$

In the SC method we were looking for patterns across units that were stable over time with the goal of constructing a synthetic control unit that is a weighted average of the control unit rows. In contrast to this, the MC estimator attempts to estimate an entirely new matrix $\mathbf{L}$ that is a low rank approximation of $\mathbf{M}$. A matrix with lower than full rank has columns and rows that can be constructed as linear combinations of the other columns and rows. Hence, the MC estimator searches for patterns both across units and time in order to find a simplified version of the matrix of interest, which can then be used to make predictions of the unknown values denoted by question marks.

The SC method can, in this exposition, be framed as a vertical regression where the control unit rows are used to estimate the treated unit row. It is possible to consider a flip side approach where the treated time periods are regressed on the pre-treatment periods, which would amount to a horizontal regression with the outcome being the vector of post-treatment periods. As pointed out in Athey et al. (2021), the MC estimator can be thought of as nesting these approaches by looking at both types of patterns (across units and time periods) simultaneously, such that it is not necessary to make a choice choose between the two.

5.4.2 Theory

The method proposed in Athey et al. (2021) frames the estimation of the unobserved potential outcomes in the absence of treatment as a matrix completion problem. Using the observed untreated outcomes of the control units and the pre-treatment outcomes of the treated unit, the method imputes values of the "missing" counterfactual outcomes for the treated unit (corresponding to $Y_{1t}^N, t > T_0$ in Rubin's Model for Causal Inference). Intuitively, this estimator uses more of the information in the data to estimate the treatment effect compared to the SC method, utilizing not only stable patterns across units but also patterns across time periods per se.

Here we present a version of the estimator that allows for individual covariates as well as time and unit fixed effects. We will not be using unit fixed effects in the actual estimation as our series are already transformed to be centered at zero, but we keep it here for completeness. The setup of the problem is such that we observe N units over T periods. For each combination of unit $j \in \{1, 2, ..., N\}$ and period $t \in \{1, 2, ..., T\}$, the unit can either be exposed to a binary treatment or not, with $W_{jt} = 1$ indicating that the unit is treated and $W_{jt} = 0$ otherwise. We observe the $N \times T$ matrix $\mathbf{Y}$ of realized outcomes, the $N \times T$ treatment indicator matrix $\mathbf{W}$, and $H \times 1$ vectors of unit/time specific variables V_{jt}. Furthermore, we define $\mathcal{M}$ and $\mathcal{O}$ as the sets of pairs of indices (j, t) corresponding to missing elements with $W_{jt} = 1$ and observed elements $W_{jt} = 0$ respectively. The focus of the method is to retrieve the missing elements in the matrix $\mathbf{Y}(0)$, that contains realized outcomes for $(j, t) \in \mathcal{O}$ and missing entries for $(j, t) \in \mathcal{M}$, to be able to estimate the treatment effect. Given this setup, Athey et al. (2021) models

the outcome variable matrix of interest, $\mathbf{Y}(0)$, as

$$\mathbf{Y}(0) = \mathbf{L} + \Gamma \mathbf{1}_T^\top + \mathbf{1}_N \Delta^\top + V_{jt}^\top \beta + \boldsymbol{\varepsilon} \quad \text{where } \mathbb{E}[\boldsymbol{\varepsilon} | \mathbf{L}, \Gamma, \Delta, V_{jt}] = 0, \tag{6}$$

where the elements of $\boldsymbol{\varepsilon}$ are σ-sub-Gaussian and can be thought of as measurement errors. Our goal is to estimate $\mathbf{L}, \Gamma, \Delta$ and β. First, we follow Athey et al. (2021) and define the operator $\mathbf{P}_{\mathcal{O}}(\mathbf{A})_{jt}$ as

$$\mathbf{P}_{\mathcal{O}}(\mathbf{A})_{jt} = \begin{cases} A_{jt} & \text{if } (j,t) \in \mathcal{O}, \\ 0 & \text{if } (j,t) \notin \mathcal{O}, \end{cases} \quad \text{and} \quad \mathbf{P}^{\perp}_{\mathcal{O}}(\mathbf{A})_{jt} = \begin{cases} 0 & \text{if } (j,t) \in \mathcal{O}, \\ A_{jt} & \text{if } (j,t) \notin \mathcal{O}, \end{cases} \tag{7}$$

for any matrix $\mathbf{A}$. Intuitively, this operator works like a mask that sets the values of the input matrix that are not observed, i.e. where $(j,t) \notin \mathcal{O}$, to zero. The proposed estimator is

$$(\hat{\mathbf{L}}, \hat{\Gamma}, \hat{\Delta}, \hat{\beta}) = \arg \min_{\mathbf{L}, \Gamma, \Delta, \beta} \left\{ \frac{1}{|\mathcal{O}|} ||\mathbf{P}_{\mathcal{O}}(\mathbf{Y} - \mathbf{L} - \Gamma \mathbf{1}_T^\top - \mathbf{1}_N \Delta^\top - V_{jt}^\top \beta)||_F^2 + \lambda ||\mathbf{L}||_* \right\}, \tag{8}$$

where $||\cdot||_F$ is the Fröbenius Norm[15] and $||\cdot||_*$ is the Nuclear Norm[16]. The regularization (or penalty) term $\lambda ||\mathbf{L}||_*$ has the purpose of rewarding solutions to the optimization problem, $\hat{\mathbf{L}}$, that has low ranks. Intuitively, attaching a positive coefficient to the sum of singular values of the matrix $\mathbf{L}$ in the objective function will punish high-rank solutions since the number of non-zero singular values of a matrix is equal to its rank. The optimal value of the penalty parameter, λ, is chosen through a cross-validation process. Specifically, we draw K random subsets $\mathcal{O}_k$ from the set of observed data $\mathcal{O}$, where the size of these subsets is chosen such that the fraction of elements that are in the subsets match the fraction of non-missing values in the entire data set. In other words, the chosen size of $\mathcal{O}_k$ should ensure that $|\mathcal{O}_k|/|\mathcal{O}| \approx |\mathcal{O}|/(N \cdot T)$.[17] Then, for some chosen sequence of candidate penalty factors $\lambda_1 > \cdots > \lambda_L = 0$ with a large enough λ_1, we estimate the model on each subset $\mathcal{O}_k$. We then evaluate for each of the estimators the average squared error on $\mathcal{O} \setminus \mathcal{O}_k$, i.e. the artificially missing values that are produced by estimating the model only on $\mathcal{O}_k$. The optimal λ is the candidate that minimizes the averaged squared error over the K validation sets.

[15] $||\mathbf{L}||_F = \left(\sum_{j \in [N]} \sigma_j(\mathbf{L})^2 \right)^{1/2}$ where $\sigma_j(\mathbf{L})$ are the ordered diagonal elements of $\mathbf{\Sigma}$ in a Singular Value Decomposition of $\mathbf{L}$.

[16] $||\mathbf{L}||_* = \sum_{j \in [N]} \sigma_j(\mathbf{L})$

[17] Note here that K has a slightly different interpretation than in standard k-fold cross-validation, where k refers to the number of equally sized subsets into which the training data is split. In our case, K does not determine the relative size of the training and validation sets, as this is forced to be equal to the fraction of missing values in the entire data set, but merely the number of times this procedure is repeated.

5.4.3 Analysis and results

In this section, we apply the MC method to estimate the effect of the 2019 smoking ban. In our main specification, presented here, we control for the unit-time-varying covariates temperature and cigarette prices. The reason for using price instead of tax rates in this analysis is that there is little variation in cigarette tax rates on a monthly basis since most tax changes occur annually or semi-annually at most. Temperature is defined as the mean temperature in a given country during a given month. Price is defined as the value of cigarette sales (denoted in EUR), obtained from the Nielsen data, divided by number of sold packs. Note however that this is an imperfect measure of price, as it also reflects changes in the composition of cigarette sales (this measure would fall if people started buying cheaper cigarettes).[18] Finally, we include time fixed effects in the model. We do not include unit fixed effects since the series are already centered at zero because of our transformation of the outcome variable into percentage deviations from the pre-treatment mean. Further, we use the default value of $K = 5$ in the cross-validation procedure used to determine the penalty term.

In Figure 20, we plot the actual and estimated outcomes for Sweden. Compared to Figure 12, the fit in the pre-treatment period is significantly better. This is likely a result of overfitting on part of the MC algorithm, which is more flexible than the SC method, so we should be careful in interpreting this as evidence for the model providing good predictions of the counterfactual outcomes. Hence, it is particularly important to rely on the placebo test results provided later in this section. We also note that the counterfactual outcomes in the post-treatment periods consistently lie slightly below the observed outcomes. While at first glance this may seem like a surprising result given our hypothesized negative treatment effect, we must compare the size of this deviation to the random deviations produced in our country placebo test before drawing any conclusions about the significance of this result. In Figure 21, we plot the gaps between the actual and estimated outcomes, noticing again that these are substantially smaller in size compared to the deviations produced in the SC model. However, consistent across the three models presented thus far is that the largest discrepancy between the actual and estimated data in the pre-treatment periods is found in the peak of the summer of 2018, where Sweden's cigarette sales deviations differ the most from those of the donor pool countries.

At this point, an advantage that the SC method has over the MC method becomes evident, namely that that the latter lacks the transparency of the former. Notably, we cannot present anything similar to the country weights in Table 2 or the predictor weights in Table 4 since the MC estimates are not based on a weighted average of the control countries' outcomes. We can, however, provide some insight into to what extent the choice of λ influences the estimated counterfactual outcomes, which we will come back to later in this section. Further, Figure A7 in the Appendix shows the impact on the estimated counterfactuals of excluding any of the covariates as well as the time fixed

[18] However, we see this as the best approach available to us given the difficulties in finding price indices for cigarettes that are homogeneously defined and measured over the countries in our sample.

Figure 20
Actual and Synthetic Sweden (MC)

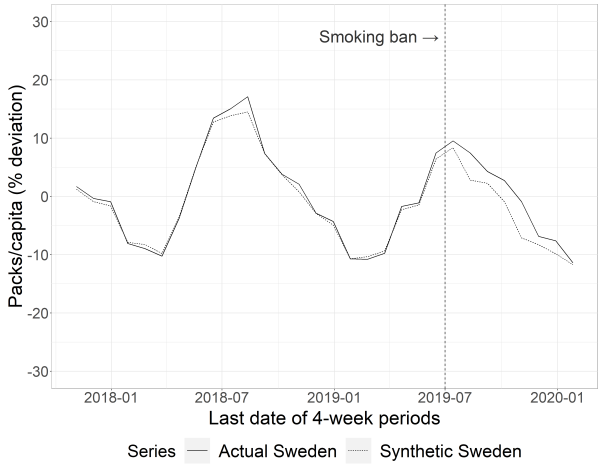
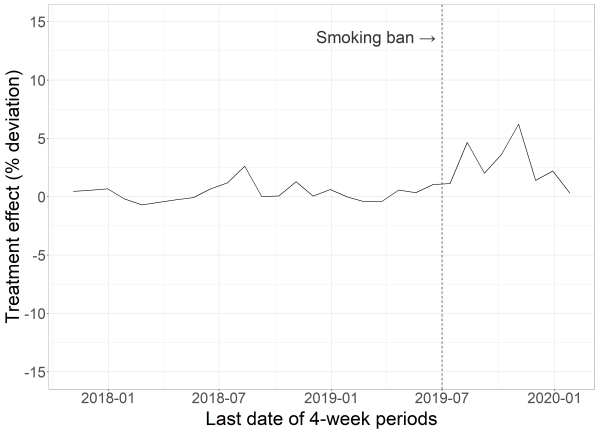


Figure 21
Treatment effect (MC)



effects, which is very small.

Next, we perform the same robustness checks as we did for the SC model. In Figure 22, we present the results from the country placebo tests. Figure 23 and Figure 24 show the in-time and in-time/country placebo tests respectively and in Figure 25 the leave-one-out analysis is presented. The country placebo test gives similar results as in the SC model, with the placebo treatment effects ranging in the interval from -5 to 5 percentage points, suggesting that what seemed like an indication of a positive treatment effect in Figure 20 is not too different in magnitude from the estimates of the placebo effects. The in-time placebo test shows that the model is able to track the post-placebo treatment outcomes well. The leave-one-out analysis shows that the MC estimates are much less sensitive to the inclusion or exclusion of certain donor pool countries than the SC model, both in the pre-treatment and post-treatment periods. This lends further support to the observation made in relation to the relaxed SC specification; when allowing for a more flexible model specification, the large importance of Latvia that we saw in the default SC specification diminishes dramatically, making results across different restricted donor pools very similar.

Overall, the performance of the MC estimator seems to be rather similar to that of the SC method, albeit with better pre-treatment fit and less reliance on certain control units. The model's good performance on all of the placebo tests suggests that the close pre-treatment fit is not entirely driven by overfitting. As mentioned, while the results in this section point towards treatment effect estimates that are in fact slightly more positive than what is found with the SC method, this effect does not appear extreme in comparison with the prediction errors generated by the country placebo tests. Hence, we see the results of this section as further evidence towards the conclusion that the smoking ban of 2019 did not have any effect in the short run.

However, we note that the cross-validation procedure used here to select a value for

Figure 22
Country placebo tests (MC)

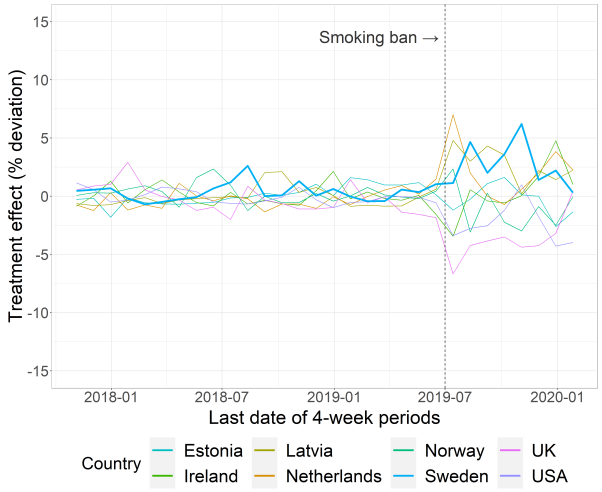


Figure 23
In-time placebo test (MC)

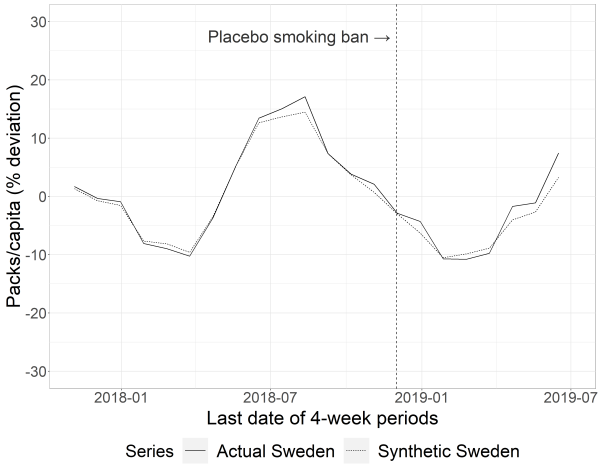


Figure 24
In-time country placebo tests (MC)

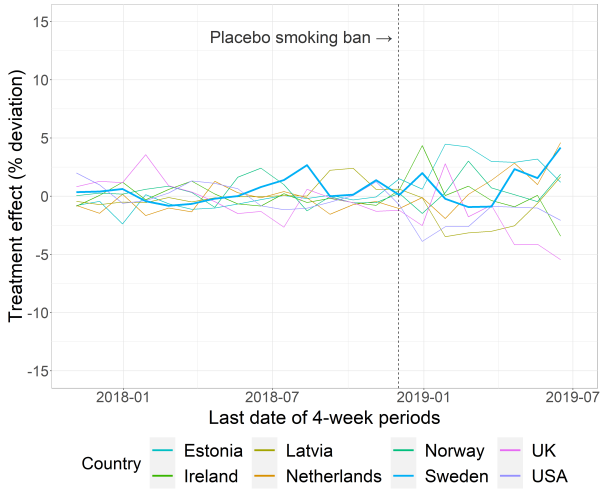
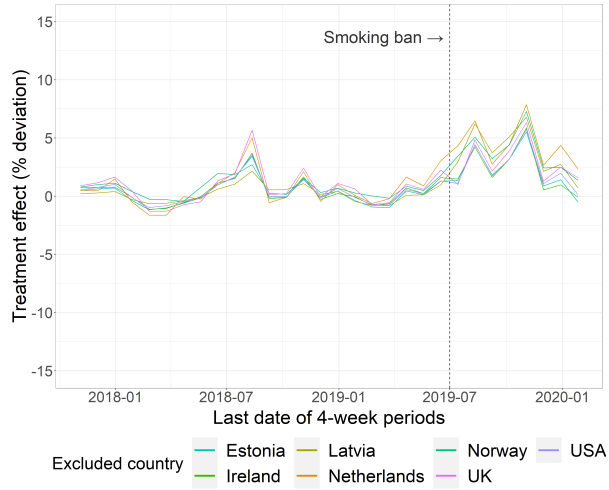


Figure 25
Leave-one-out analysis (MC)



the penalty term λ has one potential shortcoming. In particular, the MSPE that the procedure is set to minimize when choosing the value of λ is based on validation sets that are randomly drawn from the entire observed data matrix. However, in a panel data setting, we would care more about the ability of the model to accurately predict the observations that are close to being treated. In our setting, close to being treated means two things: being an observation belonging to Sweden, as well as being close to the treatment date. An approach taking this into account is to mimic the cross-validation procedure that was used for the SC method and select the value of λ that minimizes the MSPE of Sweden's cigarette sales over a set of pre-treatment periods, which we will do next. This ensures that λ is chosen so as to achieve the degree of penalization that

maximizes the ability to predict Sweden's cigarette sales over a set of periods close to the treatment date, and not those of other donor pool countries.

We detail how this alternative cross-validation procedure takes place, step-by-step, since it illuminates the way in which the MC method works. In Figure 26, we run the MC algorithm using the same placebo treatment date of December 1, 2018, as in the in-time placebo analysis above, letting Sweden be the treated unit. However, we now vary the λ parameter (penalty term) between 0 and 10 by steps of 0.1. Figure 27 plots the MSPE over the post-placebo treatment period against the values of λ. Note that the λ parameter is not independent of the scale of the data, so the range over which to search cannot be determined a priori. What is important is that the range covers a plausible minimum, which seems to be the case here. Note that when there is no regularization at all, i.e. when $\lambda = 0$, the pre-treatment fit is perfect since the estimator will by design propose the known values Y_{it} for all $(i, t) \in \mathcal{O}$, i.e. all observed data points. This highlights the crucial part that the penalty term plays in the MC estimator. Overall, the post-placebo treatment fit does not vary that much over the range of penalty term values proposed here, but this is to a large extent an artifact of the covariates and time fixed effects doing a bit of work in the predictions. Without these, the fluctuations are much larger especially for values of λ close to zero (see Figure A8).

Figure 26
Sensitivity of in-time placebos to λ

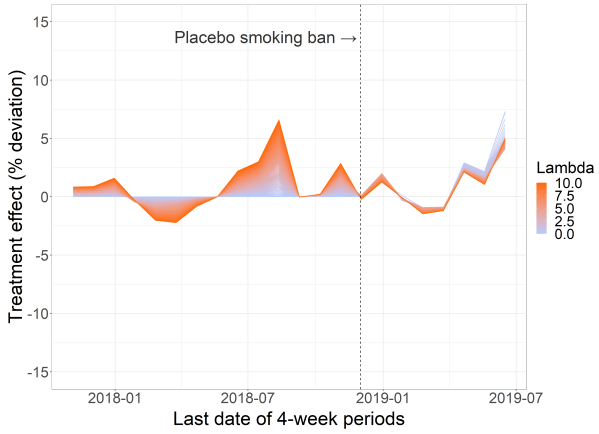
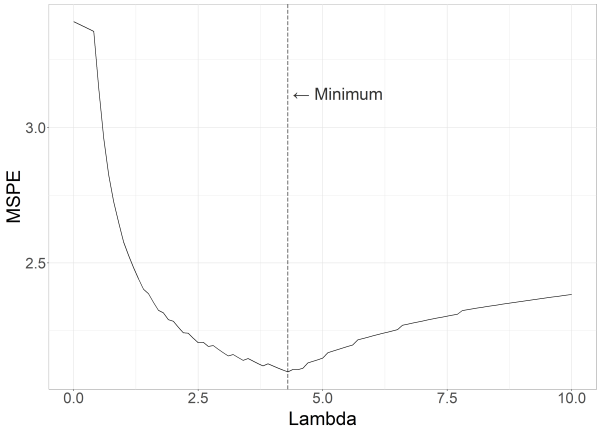


Figure 27
Relation between MSPE and choice of λ



When using the approach described above for choosing the value of λ—minimizing the MSPE of Sweden's cigarette sales over a set of pre-treatment periods—we get the solution that $\lambda = 4.3$ indicated by the dashed line in Figure 27, so we keep this as our chosen penalty parameter and conduct the same main estimations as above. The results are shown in Figure 28 and 29. The fit as well as the estimated treatment effect are almost identical to the previous results, indicating that the chosen value of λ seems rather robust in our case.

As a final investigation into the impact of λ on our results, and the ability of the MC estimator to find a low-rank approximation of the outcome matrix, we now turn to

Figure 28
Actual and Synthetic Sweden (Custom λ)

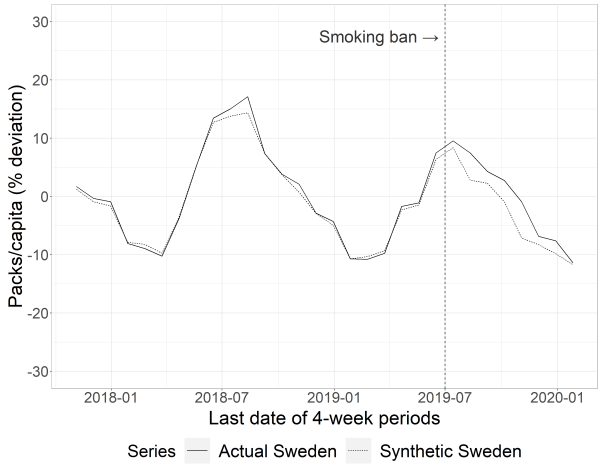
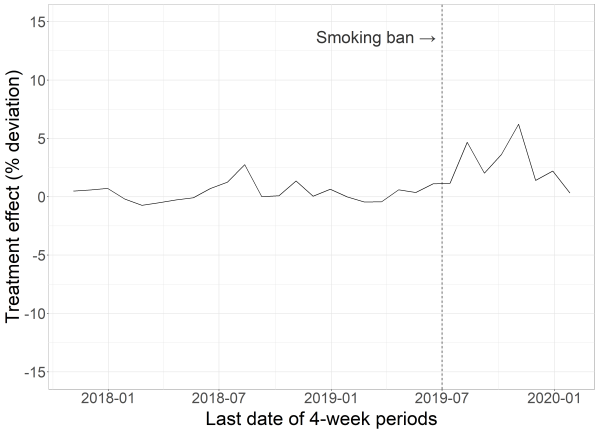


Figure 29
Treatment effect (Custom λ)



look explicitly at the rank of the estimated $\mathbf{L}$ matrix. Using the same specification as for the main results, i.e. including time fixed effects as well as price and temperature as covariates, we now compute the rank of the estimated $\mathbf{L}$ matrix when using values of λ in a range from 0 to as high as 30 by steps of 0.1.[19] This is shown in Figure 30. Interestingly, for the value of λ chosen by cross-validation, 4.3, the rank is only reduced to 7 from a maximum of 8. Hence, reducing the rank further does not seem to lead to a greater ability to predict the counterfactual outcomes for Sweden from a cross-validation perspective. This could be due to a lack of low-rank relationships in the raw data, which would not be too surprising given the relatively small dimensions of our outcome matrix. Further, it could also be that the time fixed effects and covariates are doing the main work in the predictions, such that the marginal gain in predictive ability of reducing the rank of the $\mathbf{L}$ matrix is rather negligible.

As λ increases, the algorithm attaches greater importance to reducing the rank of the estimated $\mathbf{L}$ matrix, which mechanically causes it to choose a solution with a lower rank. What is more interesting is to see what impact this has on the estimated main treatment effect. This relationship is shown in Figure 31. As we can see, increasing the λ parameter all the way up to 30, effectively leading to a solution in which $rank(\hat{\mathbf{L}}) = 2$, does not seem to have a dramatic impact on the results. In particular, the effect is always positive and does not appear that extreme in relation to the country placebo tests presented earlier in this section. This leads us to believe that the previous results using the cross-validation methods of selecting λ are quite robust, even though it leads to a solution with a relatively high rank.

In this section, we have seen that the MC method produces results that point in the same direction as the SC results, namely that there is no indication of a treatment effect of the Swedish smoking ban in the expected direction. We have also investigated

[19]The rank is computed using QR decomposition in R.

Figure 30
Relation between rank of $\hat{\mathbf{L}}$ and λ

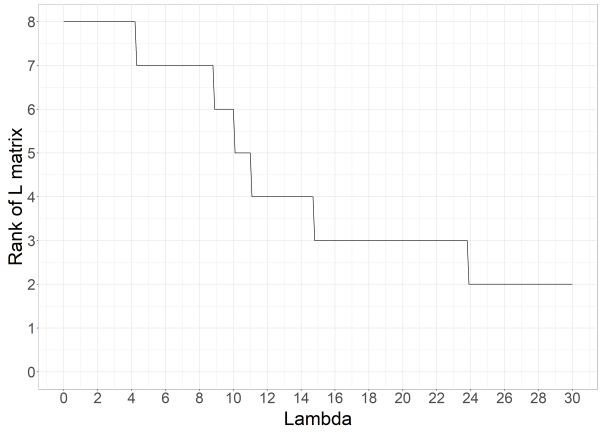
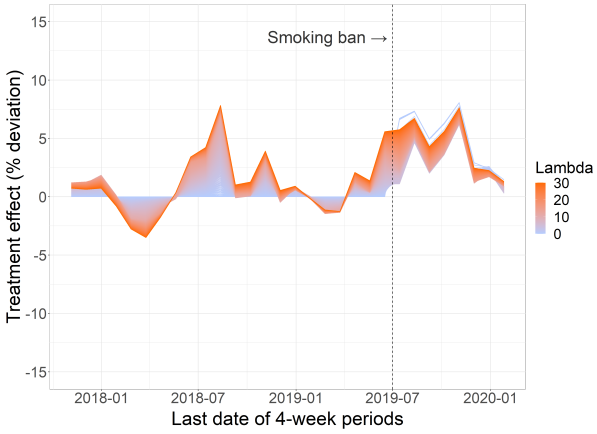


Figure 31
Sensitivity of main treatment effect to λ



the impact of the penalty term λ, the fixed effects and the covariates on the results to be transparent about the workings of the method, although it is fundamentally not as transparent as the SC estimator. In doing so, we have also attempted to provide an intuitive understanding of this relatively new methodological tool. This concludes our description of the main results of our study and, next, we turn to a number of robustness checks and extensions.

6 Robustness checks and extensions

In this section, we consider the robustness of our main results. First, we investigate the impact of our transformation of the outcome variable on our results, by re-estimating our main specification (the default SC method) on cigarette sales per capita in levels as well as after simply demeaning the series, and then comparing the obtained results to our main results. We also turn to data on sales of a substitute to cigarettes—snus—and the number of calls to a smoking quit line, to examine whether we see any indication of an effect of the Swedish smoking ban of 2019 in these related outcomes. We end this section with a discussion about the power of our study.

6.1 Transformation of the outcome variable

In the main analysis of this paper, we transform the outcome variable of cigarette sales per capita into percentage deviations from the pre-treatment mean. As discussed, there are several motivations for this. First, there are few reasons to have the SC and MC methods matching on level of sales, since this also reflects the market coverage of Nielsen in the various countries. Hence, even if we would consider it important to have the synthetic control unit resemble Sweden in the *true* levels of sales, this would not be achieved by matching on levels in our data due to the varying degree of market coverage that it

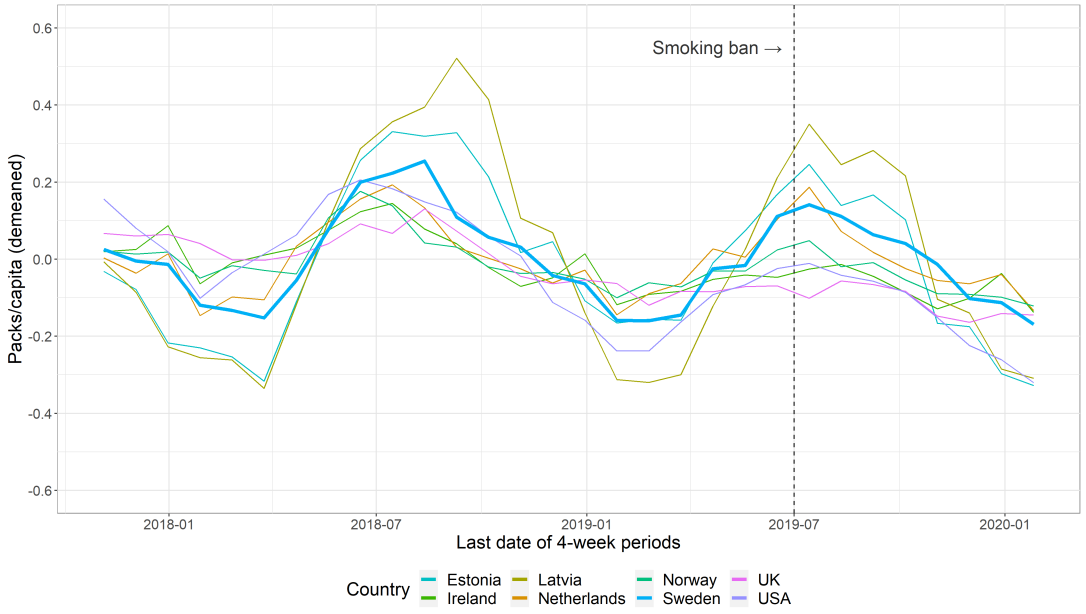
reflects. Second, simply demeaning the data creates larger fluctuations for countries that have a high level of cigarette sales, since these also have a greater degree of seasonality in absolute terms, but less so in percentages of baseline consumption.

However, in this section we will first show results when the series are only demeaned, and not transformed into percentage deviations. This transformation can be likened with including an intercept in a DID analysis, which is standard practise in the literature, and it is proposed in Abadie (2020) as a solution in situations where it is inappropriate or undesirable to match on levels. We will show this for the standard SC method to see whether the results differ significantly from the main analysis, in which case the robustness of our results may be seen as questionable. Further, in the last part of this section, we will also discuss results when the data are not transformed whatsoever and hence analysed in pure levels. In both of these cases, the results are very much in line with the main results of this paper.

Figure 32 shows the monthly times series of cigarette sales per capita when demeaned with respect to the pre-treatment means by country. As we can see in this figure, the fluctuations of Estonia and Latvia are quite extreme compared to the other countries. Hence, it is likely not possible to obtain good estimates of Estonia and Latvia in a country placebo analysis. However, it may well be possible to estimate the counterfactual outcomes of Sweden. Note that since the data are only demeaned here, the unit of measurement is cigarette packs sold per capita and not percentage deviations from the pre-treatment mean as in previous sections.

Figure 32

Monthly time series of cigarette packs sold per capita by country, demeaned



Source: Authors' rendering of data from Nielsen.

In Figure 33 and 34, we show the outcomes for actual and synthetic Sweden as well as the treatment gap. The pre-treatment fit is surprisingly good and the ability of synthetic Sweden to capture the peak during the summer of 2018 is better than in the main analysis, presumably since there now exists donor pool units in Estonia and Latvia that are more extreme relative to Sweden. We see no sign of a treatment effect in the anticipated direction following the treatment date, but a slight positive effect towards the late autumn and winter. This effect peaks at 0.1 packs per capita, which amounts to roughly 6.5% of the pre-treatment mean of Sweden. Hence, the direction and magnitude of these results are similar to those in the main analysis. Tables 6-8 show the characteristics of the synthetic control unit and its matching on observables. As we can see, Latvia receives a smaller weight than in the main analysis, essentially switching places with Estonia. Compared to the main analysis, the predictor balance is rather similar except for a worse match on the mean growth in cigarette sales between 2013 and 2018. In particular, the synthetic control unit is on a more negative long-term trend in cigarette sales than actual Sweden. The difference amounts to around 1 percentage point in the long-term trend of annual cigarette sales, so that the estimated synthetic unit may be slightly biased downwards. However, the magnitude of this possible bias is likely quite small.

Figure 33
Actual and Synthetic Sweden (SC, demeaned)

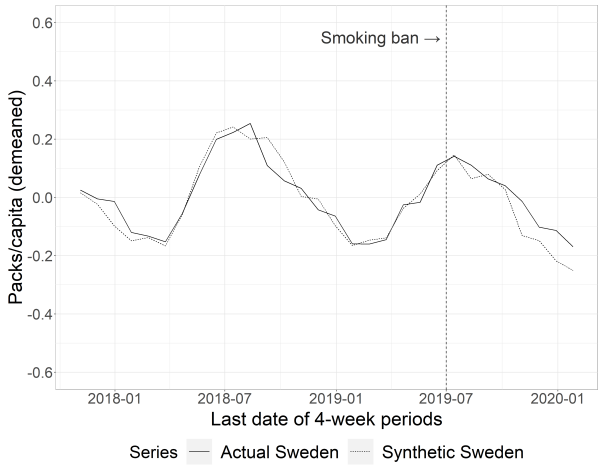


Figure 34
Treatment effect (SC, demeaned)

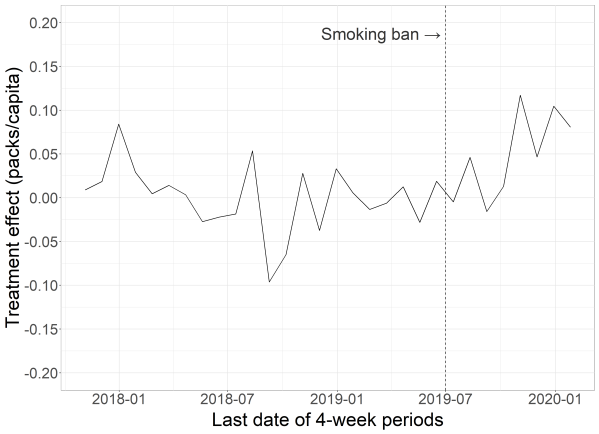


Table 6
Country weights (SC, demeaned)

	W
Estonia	0.411
Norway	0.348
USA	0.144
Latvia	0.079
Netherlands	0.015
UK	0.002
Ireland	0.001

Table 7
Predictor weights (SC, demeaned)

	V
Seasonal temperature difference, 2006–2016 (°C)	0.055
Mean growth in cigarette sales, 2013–2018 (%)	0.002
Prevalence of occasional smoking, 2018 (%)	0.563
Mean growth in tax rate, 2017–2019 (%)	0.090
Outcome mean, Jun–Aug 2018	0.001
Outcome mean, Dec–Feb 2018–19	0.136
Outcome mean, Apr–Jun 2019	0.154

Table 8
Predictor balance (SC, demeaned)

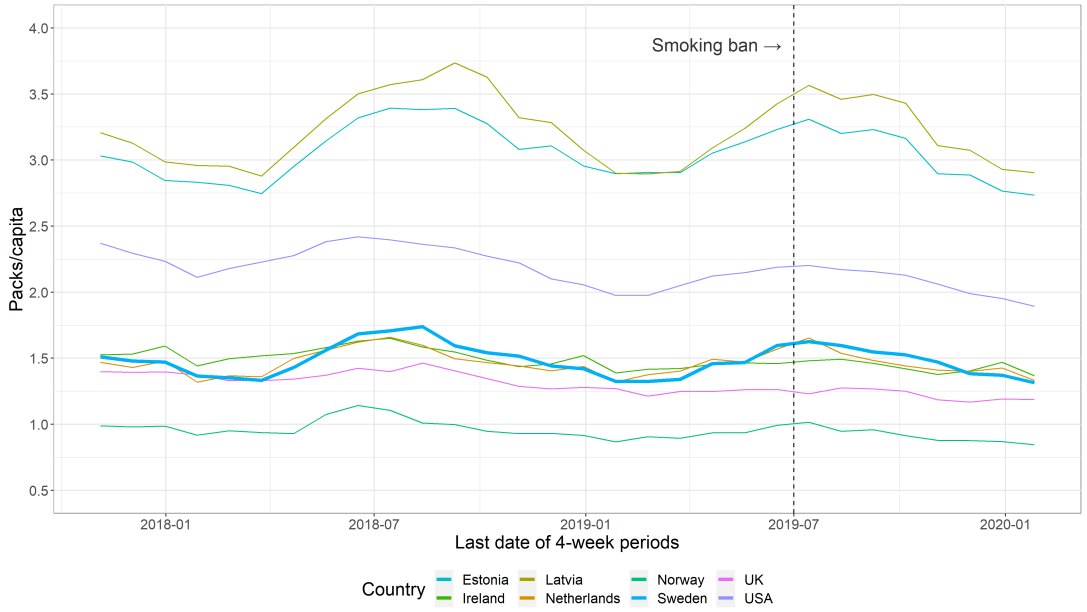
	Treated	Synthetic	Sample mean
Seasonal temperature difference, 2006–2016 (°C)	23.447	22.589	18.458
Mean growth in cigarette sales, 2013–2018 (%)	-1.441	-2.672	-3.142
Prevalence of occasional smoking, 2018 (%)	4.800	4.812	3.814
Mean growth in tax rate, 2017–2019 (%)	1.427	1.517	2.575
Outcome mean, Jun–Aug 2018	0.226	0.221	0.188
Outcome mean, Dec–Feb 2018–19	-0.088	-0.089	-0.089
Outcome mean, Apr–Jun 2019	0.023	0.022	-0.005

Source: Data on seasonal temperature difference 2006-2016 is collected from World Bank (2021a). Mean growth in cigarette sales 2013-2018 is calculated using data from European Commission (2020). Data on prevalence of occasional smoking 2018 is collected from WHO (2021). Mean growth in tax rate 2017-2019 is calculated using data from European Commission (2021). Outcome means Jun-Aug 2018, Dec-Feb 2018-19 and Apr-Jun 2019 are based on data from Nielsen.

Before concluding, we briefly discuss the results from the SC estimator when we do not transform the cigarette sales per capita at all. This amounts to an analysis of cigarette sales in actual levels, shown in Figure 35. It should be noted that in this case, we need to reconsider our choice of predictor variables. Previously, we have disregarded variables that are useful in predicting the level of cigarette sales and instead opted for factors that can explain short-term trend patterns and seasonal fluctuations. In this specification, however, this rationale is no longer appropriate. Hence, we opt for a set of predictors that are (by intention) very similar to those used in Abadie, Diamond, and Hainmueller (2010) that studies cigarette sales in California, following a new law aimed to reduce smoking, using the SC method and data in levels. In particular, we include the mean price of cigarettes over the pre-treatment period as well as the share of the population aged 15-24 (OECD 2021a), mean alcohol consumption per capita (OECD 2021b) and log GDP per capita (OECD 2021c) over the period of 2016-2018. In addition to this, we include the three lagged outcome means as before but now in levels of cigarette sales. Figure 36 and 37 show the outcomes for actual and synthetic Sweden and the treatment gap. The patterns here are very similar to those in the main results section, with a

deviation during the summer of 2018 and a treatment effect in the post-treatment period that is slightly positive but close to zero. Given the pre-treatment mean of cigarette sales per capita in Sweden, that is roughly equal to 1.5, the positive deviations correspond to 3-4% which is similar to what we have seen throughout this paper.

Figure 35
Monthly time series of cigarette packs sold per capita by country, in levels



Source: Authors' rendering of data from Nielsen.

Figure 36
Actual and Synthetic Sweden (SC, in levels)

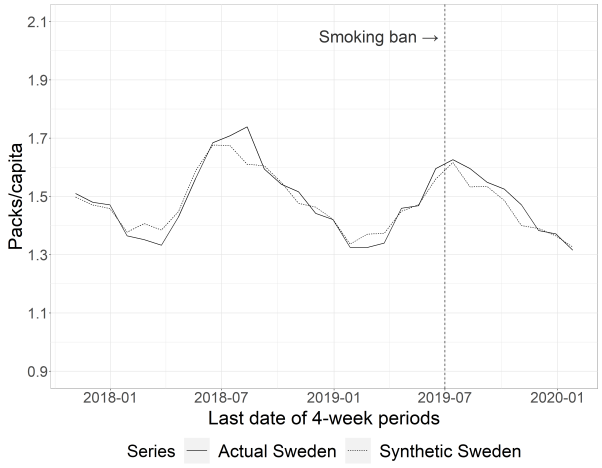
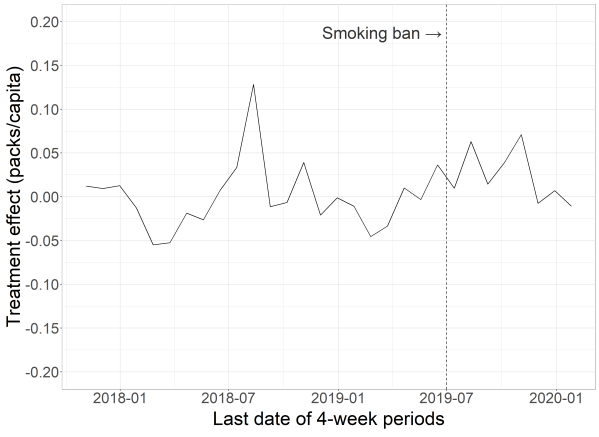


Figure 37
Treatment effect (SC, in levels)



Tables 9 and 10 show the country and predictor weights underlying synthetic Sweden. In this specification, Norway receives a very large weight while Netherlands and Latvia both range around 0.2. This is not surprising given what we have seen earlier—the

seasonal patterns in Latvia and the Netherlands appear to match that of Sweden quite well, but some weight must be attached to a country with a lower level of cigarette sales as well given the disparity in levels between Sweden and Latvia, which is likely why Norway also receives a significant weight. What might be slightly surprising here is that the predictors—that, again, are very similar to those used in Abadie, Diamond, and Hainmueller (2010)—appear to have little predictive power in our application. However, given the aforementioned reason to not match on levels due to the varying market coverage of the Nielsen data, this is not a strange result. Even if there exists a clear relationship between e.g. the level of GDP or cigarette prices with the true level of cigarette sales, this link is clearly weakened when sales are measured differently across countries. The fact that we observe such small predictor weights thus strengthens our belief that some transformation of the outcome variable is appropriate in our application. In terms of predictor balance, shown in Table 11, the low weight associated with all predictors except for the lagged outcomes naturally results in a rather poor matching overall, albeit not as bad as one might think. What stands out is that the cigarette price of synthetic Sweden is substantially greater than actual Sweden, which is driven by Norway having

Table 9
Country weights (SC, in levels)

	W
Norway	0.601
Netherlands	0.211
Latvia	0.180
USA	0.004
Ireland	0.003
Estonia	0
UK	0

Table 10
Predictor weights (SC, in levels)

	V
Retail price, 2017–2019 (EUR)	0
Alcohol consumption, 2016-2018 (Litres/capita)	0.003
Percent aged 15-24, 2016-2018 (%)	0
Ln(GDP/capita), 2016-2018	0.002
Outcome mean, Jun–Aug 2018	0.135
Outcome mean, Dec–Feb 2018–19	0.590
Outcome mean, Apr–Jun 2019	0.270

Table 11
Predictor balance (SC, in levels)

	Treated	Synthetic	Sample mean
Retail price, 2017–2019 (EUR)	5.761	10.008	8.005
Alcohol consumption, 2016-2018 (Litres/capita)	7.133	7.620	9.467
Percent aged 15-24, 2016-2018 (%)	11.607	12.005	11.689
Ln(GDP/capita), 2016-2018	10.858	10.890	10.811
Outcome mean, Jun–Aug 2018	1.711	1.654	2.155
Outcome mean, Dec–Feb 2018–19	1.396	1.407	1.877
Outcome mean, Apr–Jun 2019	1.508	1.494	1.961

Note: Retail price refers to the price of a cigarette pack and is averaged over the pre-treatment period, and hence not over the entirety of 2017 and 2019. Source: Retail price of cigarettes 2017-2019 is calculated based on data from Nielsen. Data on alcohol consumption 2016-2018 is collected from OECD (2021b) and percent aged 15-24 2016-2018 from OECD (2021a). Ln(GDP/capita) 2016-2018 is calculated using data from OECD (2021c). Outcomes means Jun-Aug 2018, Dec-Feb 2018-19 and Apr-Jun 2019 is based on data from Nielsen.

cigarette prices almost twice as high as Sweden. However, given that Norway is arguably the country most similar to Sweden in the sample, we are inclined to believe that the synthetic unit resembles Sweden fairly well in other respects.

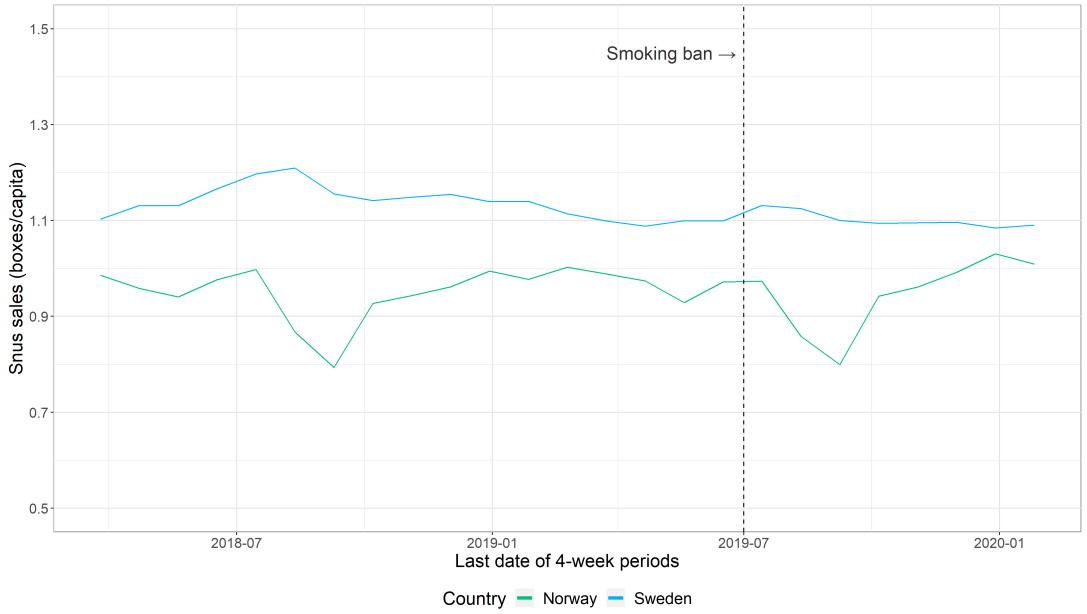
With the above analysis in mind, it seems as though the null result obtained in the main analysis is not driven by our decision to transform the outcome variable into percentage deviations from the pre-treatment mean. The differences that arise when we use demeaned data, or raw data in levels, are rather small and in particular, the direction and magnitude of the treatment effect is very similar to those observed in the main analysis.

6.2 Trends in snus sales

A rather unique feature of primarily two countries in our sample, Sweden and Norway, is that cigarettes is not the only tobacco product that is widely and frequently used. In these countries, the use of the smokeless tobacco product snus is almost on par with that of cigarettes. Seeing as both cigarettes and snus contain nicotine, it is reasonable to believe that there exists some degree of substitutability between the two products, although likely imperfectly so. In this section, we will briefly investigate what seems to have happened to snus sales around the time of the 2019 smoking ban. Given the null result obtained in the main analysis of this paper we would be surprised to see any large breaks in the trend of snus sales due to the policy, since the absence of a reduction in cigarette consumption should also exclude the possibility that a sizeable substitution to snus products took place. We have obtained data from Nielsen on snus sales in Norway and Sweden over the period from March 2018 through January 2021. However, again due to Covid-19, we restrict our analysis to periods before February 2020. This data are very similar in nature to the main data used in this paper, and the mode of aggregation (four week sums) is the same. The two time series are shown in Figure 38.

With data on two countries, a methodological option would be to perform a DID analysis. However, the absence of a clear similarity in pre-treatment trends between Norway and Sweden excludes this possibility. We will therefore restrict ourselves to a descriptive inspection of trends in this section. To begin with, we address the remarkable difference in the seasonal pattern between the Swedish and Norwegian data. Whereas the Swedish series exhibits a seasonal pattern similar to what we saw in the cigarette sales data, the Norwegian series drops between mid-July through mid-September. This difference should however not be interpreted as a cause of concern regarding the data quality, as it is likely explained by a combination of factors such as increased Norwegian cross-border shopping in summers and cultural differences in how cigarettes are substituted for snus during summer vacations. Redirecting our focus to the Swedish series at the time of the enforcement of the smoking ban of 2019, we do not see any indication of an increase in snus sales in excess of the consumption growth during summer months that is also observed in the previous year. If anything, the peak in 2019 appears to be lower than in 2018. This may however be due to a negative long-term

Figure 38
Monthly time series of snus boxes sold per capita in Norway and Sweden



Source: Authors' rendering of data from Nielsen.

trend, although it is difficult to draw any such conclusions given the short time period observed here. Overall, we cannot draw any strong conclusions based on a visual analysis of the series. However, we maintain that there is little evidence in the raw data that points toward a significant break in the snus sales series, which is in line with the results on cigarette sales found previously in this paper.

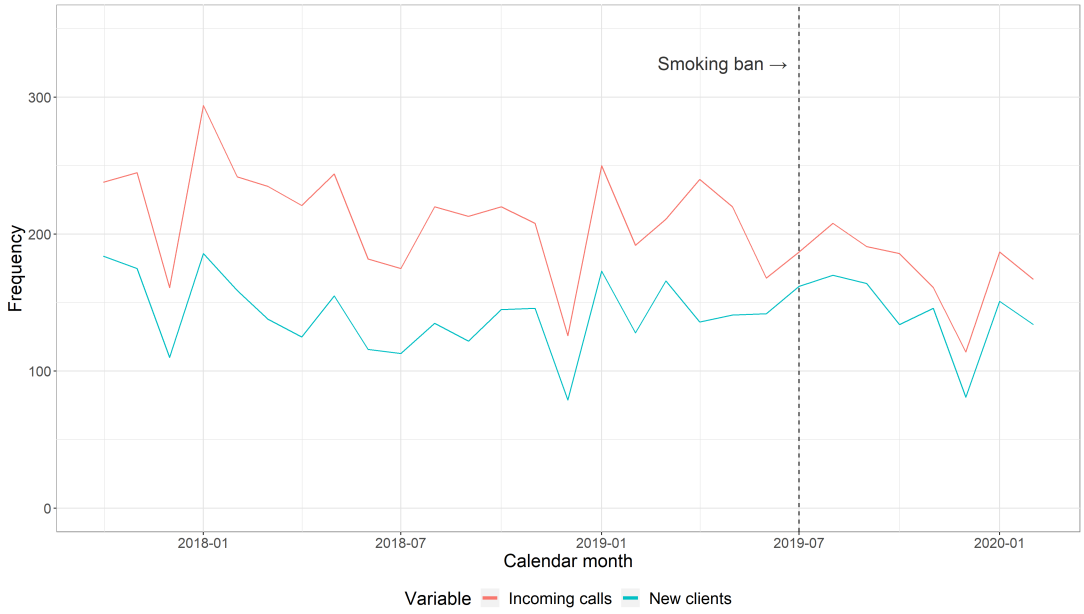
6.3 Trends in efforts to quit smoking

Even if the Swedish smoking ban of 2019 did not achieve a reduction in cigarette consumption in the short run, it may have provided smokers with incentives to make quit attempts in line with what was found in the previously discussed study of Ko (2020). Since cigarettes are highly addictive, a policy that increases quit attempts may still be successful in reducing actual smoking in the long run as smokers may require several efforts before successfully managing to quit. While performing a thorough analysis of quit attempts lies outside the scope of this paper, we devote this section to investigating trends before and after the 2019 smoking ban in data related to smoking cessation attempts.

We have acquired monthly data from October 2017 to February 2020 on the number of incoming calls to and new clients with *Sluta-Röka-Linjen* [The Quit Smoking Hotline], a free-of-charge telephone line devoted to helping people stop using tobacco products. In addition to offering support and motivation over telephone and online chats,[20] callers are able to register as clients with the hotline in order to be actively contacted and checked up

[20]The online chat was not functioning between March and August 2019, so these data are excluded from analysis.

Figure 39
Monthly time series of incoming calls and new clients at the Quit Smoking Hotline



Source: Authors' rendering of data from Sluta-röka-linjen.

on over a period of time. Both data on the number of incoming calls and newly registered clients are shown in Figure 39. Activity with the hotline peaks in January every year, presumably due to New Years' resolutions, and thereafter gradually declines for the rest of the year. A less visible but still recurring peak occurs around August/September in both 2018 and 2019, that may be explained by a tendency for people to make new attempts to live healthily after summer holidays. However, the pattern does not suggest that the peak in 2019 was higher than that of the previous year, at least for incoming calls. Newly registered clients increase more in July 2019 than in the same month the year before, but in relation to the overall volatility of the series, the difference does not appear striking. Further, there are no signs of a break in the downward and flat trends of incoming calls and new clients, respectively, following the smoking ban of 2019.

As in the case of the snus sales investigated in the previous section, we are not able to draw any conclusions about causality on basis of a simple visual inspection of pre- and post-treatment trends. However, it serves as an indication that quit attempts, as proxied by activity with the Quit Smoking Hotline, did at least not increase dramatically following the 2019 smoking ban. This is, again, in line with the results that have been presented previously in the paper.

6.4 A note on statistical power

An important consideration in all empirical research, that has received a great deal of attention among scholars in a variety of fields in recent years, is statistical power (see e.g. Ioannidis 2005; Benjamin et al. 2018). In particular, it is crucial that studies

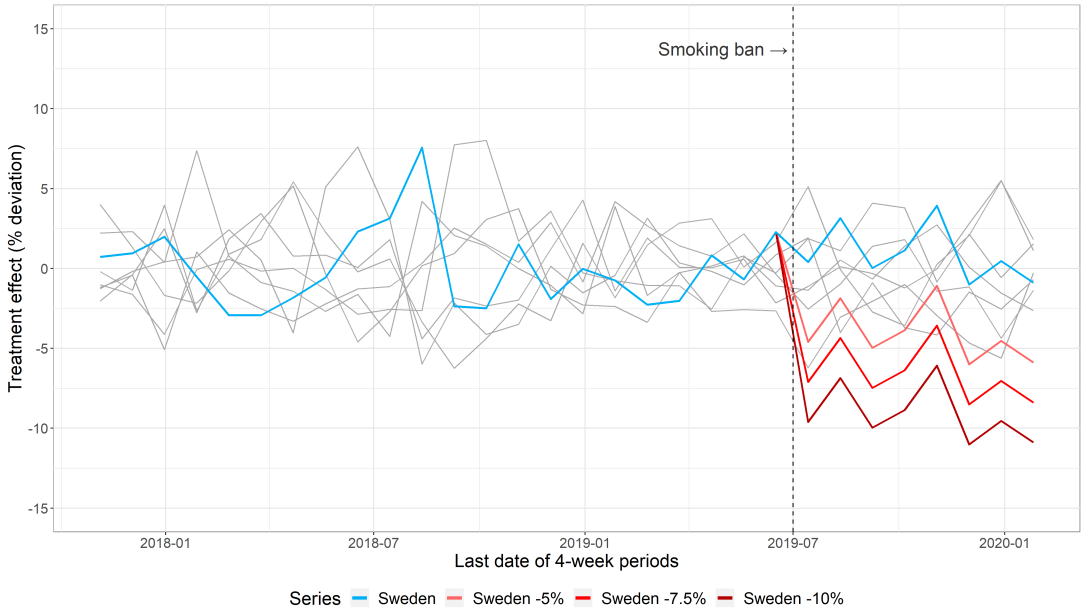
are sufficiently powered such that treatment effects of plausible magnitudes can be distinguished from noise in the data. In the context of our study, we do actually find point estimates in the opposite direction of what was expected. However, from a power perspective, it is still of great importance to discuss what conclusions could have been drawn in case we would have observed an estimated treatment effect in the anticipated direction and of a magnitude that we find economically important. Hence, we devote this last section to addressing the power of our study.

Before going into our power analysis, however, there are two points that are important to discuss. First, there are no standard procedures for formal inference or computation of statistical power for the SC method, not to mention the MC estimator, especially not for applications in which there is a limited number of donor pool units, although there is currently a lot of work being done within this area (see e.g. Firpo and Possebom 2018; Chernozhukov, Wüthrich, and Zhu 2021). Therefore, we will investigate the power of our analysis in the best way that we see possible given the lack of guidance on the topic. Second, whereas an ex ante computation of statistical power is possible and advisable for many methods, such as in the case of randomized controlled trials, it is not possible for analyses using the SC or MC methods such as ours given the current state of knowledge. Therefore, power discussions can only take place in hindsight and could therefore not have informed the decision of whether to carry out the study presented in this paper.

So far, we have relied on visual inspection of the country placebo tests to determine an interval outside of which we feel confident in excluding potential treatment effects. However, in this brief investigation we pose a fundamental question related to statistical power; how large would the treatment gap have to be for us to interpret it as evidence against the null of no treatment effect? We investigate this question by taking the estimated Swedish treatment gap from the default SC estimator, adding hypothetical treatment effects of -5, -7.5 and -10 percentage points, and plotting the three resulting series along with the results from the country placebo tests. Then, for each of the resulting post-treatment series, we evaluate whether we would have interpreted such treatment gaps as evidence against the null of no effect. The results from this exercise are shown in Figure 40.

In the case where we add a treatment effect of -5%, the Swedish treatment gap is at the edge of the bandwidth given by the country placebo tests, but still inside it in several post-treatment periods. Hence, we would consider this a borderline case in which there may be suggestive evidence against the null, but not strong enough to be able to draw any strong conclusions about the effectiveness of the policy. This is in line with what we have discussed earlier in the paper, where the country placebo test results have been interpreted as an error interval roughly bounded by the $\pm 5\%$ marks. However, when we turn to the cases of adding treatment effects on the order of -7.5% or even -10%, the Swedish treatment gap appears quite extreme relative to the country placebo tests. We maintain that in both of these cases, we would have concluded that there is considerable evidence against the null of no effect, which in our view is the strongest kind of statement that we could possibly make given our study design. In a rough sense, then, we consider

Figure 40
Imputed treatment effects in relation to country placebo test results



Note: Gray lines show placebo test outcomes for all countries in the donor pool, which are identical to those shown in Figure 14.

our study to be powered to detect effect sizes greater than 5% of the pre-treatment mean.

Compared to the effect sizes found in previous literature, an effect of more than 5% is to be seen as fairly large. Whereas effects of this magnitude have certainly been found in studies of indoor smoking bans, we would not expect this weaker type of regulation to have an effect on cigarette consumption much larger than 5%. Hence, despite the difficulty in forming priors on this effect due to the scarcity of previous evidence on outdoor smoking bans, we believe that our study is likely underpowered. However, it is worth reemphasizing that this possible lack of power does not invalidate the existence of the study. Ex ante, we had no way of addressing these power issues, and as such this could not have been a deciding factor in our choice to investigate the topic or in our design of the study.

7 Discussion

In the analyses presented above, the results all point in the same direction. In the default version of the SC estimator, the point estimates of the treatment effect are close to zero during the entire post-treatment period. However, we raised some concerns with regards to the pre-treatment fit and in particular the ability of the model to capture the cigarette sales of Sweden during the summer of 2018, that may have implications for its ability to predict post-treatment outcomes during the subsequent summer of 2019. Allowing for extrapolation, using what we refer to as a relaxed version of the SC estimator, improved the fit significantly while still producing a null result. Employing

the MC estimator yielded further improvements in terms of pre-treatment fit and, more importantly, performance on placebo tests that were good and in line with previous results from the SC estimator. Using this method, we found consistently positive point estimates of the treatment effect, in the opposite direction of our priors, but the estimates were not extreme enough to be distinguishable from a regular prediction error.

The validity of these results hinges on the ability of the SC and MC methods to produce a counterfactual unit that is a good approximation of Sweden in the absence of treatment. We believe that the placebo tests performed support this statement. Looking at the in-time placebo tests, the counterfactual units that were estimated all managed to track Sweden's cigarette sales closely during time periods prior to the treatment date, with errors amounting to only a few percent of the pre-treatment mean. Further, the prediction errors attained from the country placebo tests lie within a bandwidth of $\pm 5\%$ of the pre-treatment mean. Hence, if the true effect size of the smoking ban of 2019 would exceed around 5% of the pre-treatment mean, we would expect to see a treatment gap for Sweden that is extreme relative to the prediction errors, in terms of magnitude but perhaps also in terms of consistency; many of the prediction errors close to $\pm 5\%$ are spikes that in the subsequent periods revert back to an "effect" in the opposite direction. However, in neither of our analyses do we find any indication of such a treatment gap for Sweden.

A complementary approach to evaluating the credibility of the counterfactual units is to take a closer look at the donor pool characteristics. From our discussion in section 5.2, it is evident that Sweden stands out among the sample countries in terms of the availability of substitutes to cigarettes and the prevalence of occasional smokers relative to daily smokers. To the extent that countries differing along these parameters can be expected to respond differently to shocks, this may negatively influence the credibility of the counterfactuals. However, we do not see any theoretical reasons why this would be the case, neither is there anything in our annual or monthly data that suggests that Sweden's cigarette consumption is indeed developing differently from that of the control countries. Our examination of the country characteristics also revealed another potentially important characteristic of the donor pool, namely that Latvia and Estonia differ from the other countries, especially in smoking prevalence. At the same time, Latvia receives a large weight in our SC estimators. If the crucial assumption underlying our SC analysis—that the level of smoking is irrelevant for the response to shocks affecting seasonal fluctuations in percentage terms and the short run trend, given a certain pre-treatment trend—does not hold we could indeed be worried about the counterfactuals' heavy reliance on Latvia. However, the fact that our leave-one-out analyses suggest that the results do not change enough to alter our interpretation of them with the exclusion of Latvia and that the take-away from the level analysis of section 6.3, in which the Baltic countries are of little importance, is the same as for our main analysis indicate that the inclusion of Latvia and Estonia does not appear very problematic after all.

The legitimacy of the counterfactual units of our main estimations also rely on the aforementioned assumption in another way. Specifically, we assume that the level of

cigarette consumption in synthetic Sweden is irrelevant for its ability to predict the counterfactual outcomes of Sweden. This assumption is needed since our transformation of the outcome variable into percentage deviations from the pre-treatment mean removes the level dimension from the data. While it is difficult to test the validity of this assumption, we believe that it is supported by our robustness checks, where we show that the results do not change considerably when we instead perform the analysis on the non-transformed data. Furthermore, this is a standard assumption in the DID literature in which it is generally not considered to be problematic.

Our results allow for a varying degree of confidence in concluding that the effect of the policy was non-existent in the short-term, depending on how conservative we wish to be. First, most point estimates of the treatment effect that we find are in fact positive and not negative as expected. Since we do not believe in the true treatment effect plausibly being positive, this points towards a null effect. However, this interpretation disregards the average prediction error produced by the various models employed. These range within the $\pm 5\%$ bandwidth around the pre-treatment mean, so a more conservative interpretation would be to exclude the existence of an effect greater than 5% of the pre-treatment mean, while remaining agnostic about whether the true effect is null or simply smaller than 5%. However, relying heavily on the bandwidths produced by the country placebos has two downsides. First, the number of donor pool countries that we have is rather small, so the bandwidth cannot be interpreted as a statistical confidence interval per se but merely as an *indication* of the magnitude of a typical error. Second, since we are allowing countries in the placebo analysis to have a pre-treatment fit up to three times as bad as that of Sweden, we are in fact estimating the bandwidth on countries for which our SC/MC methods are relatively less appropriate. This leads to a possible overexaggeration of the range of the prediction error bandwidth as long as pre-treatment fit positively correlates with post-treatment fit, which is a reasonable assumption. In the end, when presented with the option we believe in the prudent rule of thumb of going for the more conservative interpretation. Hence, our interpretation of our findings is that we are able to rule out any effect of the smoking ban of 2019 on cigarette sales larger than approximately 5% of the pre-treatment mean.

To contextualize this figure, 5% of Sweden's pre-treatment mean of around 1.5 packs per capita and month is equivalent to 0.075 cigarette packs or 1.5 cigarette sticks per capita and month. Given that around 13% of the Swedish population smokes (Figure 5), this would correspond to roughly 0.5 cigarette packs or 10 cigarette sticks per smoker and month. However, as previously discussed, we should be careful in putting too much weight on these calculations since they are based on the Nielsen data in which the reported level of smoking is a function on the company's market coverage rate. While previous findings regarding smoking bans that are similar but not identical to the one studied in this paper are mixed, we regard our power to detect effect sizes greater than this as slightly low from a theoretical perspective. However, since the Swedish smoking ban of 2019 was implemented nationwide at a single point in time, researchers are naturally directed towards a cross-country analysis. This limits the potential number of control

units that can be used in order to analyze the effects of the ban, since there are only so many countries that are sufficiently similar to Sweden. This, in turn, makes analyses more powered than ours quite difficult to pursue. As a final point, however, we should note that even if our results indicate that there is little reason to believe that the smoking ban of 2019 led to any great reductions in cigarette consumption in the short-term, this does not necessarily imply that the ban was ineffective in achieving its purpose along other dimensions than that of reducing individual consumption of cigarettes. One of the primary goals of the smoking ban was to increase access to public spaces for people who are particularly sensitive to smoke. It is possible, and even likely, that the ban has been to benefit of this cause even if cigarette consumption has remained stable, if smokers simply chose to relocate their smoking temporally or spatially in response to the ban.

These results mandate a discussion on the potential explanations as to why we do not observe any substantial effect of the policy. There are essentially two possible explanations; either the smoking ban of 2019 *did* have a short-term impact on cigarette consumption, while being statistically undetectable in our analyses, or there is no such effect. Before we elaborate on each of these explanations, it should be emphasized that in none of these cases can the existence of a long-term effect be ruled out. It may be that the smoking ban of 2019 will contribute to denormalizing smoking and decreasing the general population's exposure to smokers and tobacco smoke in public spaces, leading to lower pickup rates, higher cessation rates, reduced smoking intensity or a combination of the three in the long run. Further, it could be that effect of the ban on the costs and/or benefits of smoking does not translate into changes in consumption patterns in the short-term due to the addictive nature of the cigarettes, if it takes time for people to change their smoking behavior. If this is the case, the short-term perspective of our analysis underestimates the impact of the policy. Unfortunately, it will likely prove difficult to investigate the existence of such long-term impacts due to the outbreak of the Covid-19 pandemic approximately seven months after the treatment, since the heterogeneous effects of the pandemic may well affect smoking behavior in a permanent manner. However, with the data available to us, we cannot go beyond speculations on the matter of long-term effects.

There are several potential explanations to why the short-term effect of the smoking ban of 2019 would be too small for us to detect (smaller than around 5% of the pre-treatment mean). These explanations are not mutually exclusive, and can also explain why there would be close to no effect at all. Therefore, the following discussion applies to both the case of a small effect size and that of a non-existing effect. To begin with, it may be that the increase in the costs and/or reduction of the benefits of smoking caused by the smoking ban of 2019 were not sufficiently large to impact cigarette consumption in a non-negligible fashion. In other words, the effect on the affected may be too small. For example, it could be that smokers already experienced that they had to exert significant effort to be able to smoke, and so the marginal increase in the time and effort cost of having to avoid the newly regulated areas was small. It is also conceivable that smokers responded to the ban by decreasing their visits to the regulated areas, for example meeting

friends at home where it is possible to smoke instead of at outdoor serving areas. Further, it may be that the addiction caused by cigarette consumption is strong enough so as to override the impact of marginal changes in costs or benefits of smoking, rendering the policy ineffective.

Another possible explanation is that the ban did have an impact along the lines of significantly increasing costs and reducing benefits of smoking, but only on a small group of smokers. In this case, we would get an average treatment effect close to zero in the entire population of cigarette consumers, which is what we observe in this paper. This point is related to the finding of Anger, Kvasnicka, and Siedler (2011) that smoking bans impact only the subgroups who frequently visit the affected public spaces. Specifically, translating these findings to the context of our study, it may be that too few smokers smoked a too small share of the total cigarette sales in the areas where the ban had a significant impact on the costs and/or benefits of smoking. It could also be that the smoking ban had heterogeneous effects on the cost-benefit analysis associated with smoking for different smoker types. For example, it may have had substantial impact on occasional, especially social, smokers who frequently visited the regulated environments but not on daily smokers, which would further limit the size of the affected group.

Although our analysis points towards a null effect of the smoking ban of 2019 in Sweden, there are reasons to be cautious in generalizing this result to other contexts. The effectiveness of regulations such as smoking bans presumably varies non-negligibly depending on the characteristics of the country or sub-national area that it applies to. For example, countries in which cigarette smoking is more deeply embedded in the culture by being strongly linked to social contexts outside of the home where alcohol is consumed may see larger effects of a similar policy, since it particularly increases the costs of smoking in these kinds of situations. Further, the composition of the smokers is likely also a factor determining the efficacy of the policy, at least in the short-term. Two countries with the same aggregate stock of cigarette consumption may vary along the lines of smoking intensity and age distributions, for example by having many young people smoking relatively little or, in the opposite case, a smaller group of old people that has smoked quite intensively over a long period of time. In these cases, it might be likely to expect a weaker effect of the policy in the latter case due to the buildup of heavy nicotine addiction on part of smokers. The two most unique features of Sweden that are important to consider from an external validity standpoint, is its already low level of cigarette consumption as well as the availability of a widely used and socially accepted nicotine substitute for cigarettes—snus. The impact of these features on the expected efficacy of policies like the 2019 smoking ban is, however, not straight-forward when considered jointly. Efforts to reduce smoking from an already low level by way of policy interventions may well be both harder and easier compared to doing so when smoking rates are high. It could be that the last group to give up on smoking is the one most reluctant to doing so, and that this is the segment of smokers that remains in Sweden today. On the other hand, countries in which smoking is less ingrained in the majority culture may find it easier to reduce it due to it being perceived as a less

acceptable activity. Given the large health care costs associated with smoking, there may also be a greater majority of the population exerting peer pressure on smokers to quit. Further, the availability of a close substitute in form of snus could also increase the efficacy of smoking bans by providing an alternative source of nicotine not affected by it. Taken together, it is not clear whether we should expect Sweden to be a country in which the smoking ban would be particularly effective or not.

8 Conclusion

In the bill towards the new smoking ban of 2019 in Sweden, legislators put forth two objectives of the policy: to protect people that are sensitive to or do not like secondhand smoke exposure in outdoor public venues, and to reduce smoking in society at large (Prop. 2017/18:156). This paper has analyzed the impact of the smoking ban on cigarette sales, thus evaluating the latter policy objective, restricting the analysis to cover any short-term effects that may have emerged in the seven post-treatment months before the onset of the Covid-19 pandemic in early 2020. First, we apply the SC method (Abadie 2020), in which our identification strategy is to construct a counterfactual unit from a set of untreated control countries, that is matched to Sweden on our chosen predictors as well as its cigarette sales series, measured in percentage deviations from the pre-treatment mean. We adopt both the default version of the SC estimator and a variant in which we relax the constraints on the country weights, to allow the synthetic unit to better capture the pattern in the Swedish data. Next, we extend the analysis by using the relatively new MC method (Athey et al. 2021), where the counterfactual outcomes are estimated by exploiting the correlation structure across time as well as across units in the outcome variable matrix, treating it as a matrix completion problem. In this three-step analysis, we find that the impact of the ban on the aggregate cigarette consumption in Sweden is negligible at best. While we find point estimates in the opposite direction of what is expected, these are all within the margins of error around the null. Hence, we conclude that the smoking ban was likely ineffective in achieving its objective of reduced cigarette consumption, at least in the short run. While this does not exclude the possibility of a long-term effect, emerging after one or several years due to for example denormalization of smoking, this effect will likely remain difficult to identify in statistical analyses due to the Covid-19 pandemic.

This paper contributes to literature on tobacco control policy in several ways. First, numerous previous studies have investigated the impact of smoking bans in the short-term using survey data and statistical methods such as the DID estimator. We extend this methodology by using retail scanner data, which arguably gives a more reliable measure of both the magnitude and timing of sales than self-reports on consumption, and newer statistical tools. Further, we contribute to the literature on cross-country analysis methods by providing an example of how recent developments within this literature can be useful in application. In particular, we highlight the impact of relaxing the default constraints of the SC method (Doudchenko and Imbens 2016), while also applying one of

the latest methodological advancements in the literature—the MC method (Athey et al. 2021).

The Swedish ban of 2019 is also an important piece of legislation to investigate from a tobacco control policy perspective. In most of the Western world, indoor smoking bans have been in place for at least a decade and its efficacy in combating smoking is well appreciated among policy makers. Outdoor smoking bans with a scope as wide as that of the Swedish regulation, however, is a relatively new and uncommon policy tool that allows for a great deal of flexibility when determining which spaces to include. Because of this, the scientific evidence guiding legislators in the design of such bans is scarce. By investigating the efficacy of the Swedish smoking ban of 2019 along the objective of smoking reduction, we contribute to the evidence underlying policy making decisions aimed to enhance public health around the world.

Our study is, along with Kennedy (2010) and Ko (2020), among the first to investigate the impact of outdoor smoking bans on cigarette consumption. Given the aforementioned flexibility in designing such bans, further research is needed in order to investigate their potential. Specifically, it may be that some inclusions of public venues in an outdoor smoking ban have limited impact, while others are more successful. Further study of future bans differing in their extensiveness would therefore be valuable in the process of designing new policies. Despite the weak or non-existent effects found in this paper and that of Ko (2020), there is certainly a possibility that outdoor smoking bans may be effective in the long-term. Hence, studies that extend the analysis along the time dimension are certainly needed. The span of potential policies between full-scale indoor smoking bans, and outdoor bans that are restricted only to a few public spaces, is large. Policy makers enjoy great discretion in positioning new legislation along this scale and as such, there is a dire need of more evidence on what works and what does not.

References

Abadie, Alberto (2020). "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects". *Journal of Economic Literature*, forthcoming. URL: https://www.aeaweb.org/content/file?id=12409.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program". *Journal of the American Statistical Association* 105.490, pp. 493–505.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2015). "Comparative Politics and the Synthetic Control Method". *American Journal of Political Science* 59.2, pp. 495–510.

Abadie, Alberto and Javier Gardeazabal (2003). "The Economic Costs of Conflict: A Case Study of the Basque Country". *American Economic Review* 93.1, pp. 113–132.

Anger, Silke, Michael Kvasnicka, and Thomas Siedler (2011). "One Last Puff? Public Smoking Bans and Smoking Behavior". *Journal of Health Economics* 30.3, pp. 591–601.

Athey, Susan, Mohsen Bayati, Nikolay Doudchenko, Guido W. Imbens, and Khashayar Khosravi (2021). "Matrix Completion Methods for Causal Panel Data Models". *Journal of the American Statistical Association*.

Azagba, Sunday (2015). "Effect of Smoke-Free Patio Policy of Restaurants and Bars on Exposure to Second-hand Smoke". *Preventive Medicine* 76, pp. 74–78.

Babb, Stephen, Ann Malarcher, Gillian Schauer, Kat Asman, and Ahmed Jamal (2017). "Quitting Smoking Among Adults—United States, 2000–2015". *Morbidity and Mortality Weekly Report* 65.52, pp. 1457–1464.

Bauer, Joseph E., Andrew Hyland, Qiang Li, Craig Steger, and K. Michael Cummings (2005). "A Longitudinal Assessment of the Impact of Smoke-Free Worksite Policies on Tobacco Use". *American Journal of Public Health* 95.6, pp. 1024–1029.

Benjamin, Daniel J. et al. (2018). "Redefine Statistical Significance". *Nature Human Behaviour* 2.1, pp. 6–10.

Berkeley Earth (2021). *Regional Climate Change*. Electronic dataset. URL: http://berkeleyearth.lbl.gov/country-list/ (visited on 01/24/2021).

Boes, Stefan, Joachim Marti, and Johanna C. Maclean (2015). "The Impact of Smoking Bans on Smoking and Consumer Behavior: Quasi-Experimental Evidence from Switzerland". *Health Economics* 24.11, pp. 1502–1516.

Brzozowski, Matthew, Thomas F. Crossley, and Joachim K. Winter (2017). "A Comparison of Recall and Diary Food Expenditure Data". *Food Policy* 72, pp. 53–61.

Chaloupka, Frank J. and Kenneth E. Warner (2000). "Chapter 29 The Economics of Smoking". In: vol. 1. Handbook of Health Economics. Elsevier, pp. 1539–1627.

Chandra, Siddharth and Frank J. Chaloupka (2003). "Seasonality in Cigarette Sales: Patterns and Implications for Tobacco Control". *Tobacco Control* 12.1, pp. 105–107.

Chernozhukov, Victor, Kaspar Wüthrich, and Yinchu Zhu (2021). "An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls". *Journal of the American Statistical Association*.

Cullberg, Jonas (Aug. 2, 2019). "HRF: 'Rökförbudet fungerar bra'". *Arbetet*. URL: https://arbetet.se/2019/08/02/hrf-rokforbundet-fungerar-bra/ (visited on 02/11/2021).

Doudchenko, Nikolay and Guido W. Imbens (2016). *Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis*. Working Paper 22791. National Bureau of Economic Research.

European Commission (2015a). *Smoke-free Environments*. URL: https://ec.europa.eu/health/tobacco/smoke-free_environments_en (visited on 02/16/2021).

European Commission (2015b). *Smoking of Tobacco Products by Sex, Age and Country of Birth*. Electronic dataset. URL: https://ec.europa.eu/eurostat/databrowser/view/HLTH_EHIS_SK1B__custom_563429/default/table (visited on 02/12/2021).

European Commission (2017). *Product regulation*. URL: https://ec.europa.eu/health/tobacco/products_en (visited on 02/16/2021).

European Commission (2020). *Releases for Consumption of Cigarettes 2002-2019*. Electronic dataset. URL: https://ec.europa.eu/taxation_customs/sites/taxation/files/docs/body/tobacco_products_releases-consumption.pdf (visited on 02/17/2021).

European Commission (2021). *Taxes in Europe Database v3*. Electronic dataset. URL: https://ec.europa.eu/taxation_customs/tedb/splSearchForm.html (visited on 02/09/2021).

Firpo, Sergio and Vitor Possebom (2018). "Synthetic Control Method: Inference, Sensitivity Analysis and Confidence Sets". *Journal of Causal Inference* 6.2.

Fisher, Robert J. and James E. Katz (2000). "Social-Desirability Bias and the Validity of Self-Reported Values". *Psychology & Marketing* 17.2, pp. 105–120.

Foulds, Jonathan, Lars Ramstrom, Michael Burke, and Karl Fagerström (2003). "Effect of Smokeless Tobacco (Snus) on Smoking and Public Health in Sweden". *Tobacco control* 12.4, pp. 349–359.

Frazer, Kate, Joanne E. Callinan, Jack Mchugh, Susan van Baarsel, Anna Clarke, Kirsten Doherty, and Cecily Kelleher (2016). "Legislative Smoking Bans for Reducing Harms from Secondhand Smoke Exposure, Smoking Prevalence and Tobacco Consumption". *Cochrane Database of Systematic Reviews* 2, pp. 1–191.

Gallet, Craig A. and John A. List (2003). "Cigarette Demand: A Meta-Analysis of Elasticities". *Health Economics* 12.10, pp. 821–835.

Holland, Paul W. (1986). "Statistics and Causal Inference". *Journal of the American Statistical Association* 81.396, pp. 945–960.

Hopkins, David P. et al. (2001). "Reviews of Evidence Regarding Interventions to Reduce Tobacco Use and Exposure to Environmental Tobacco Smoke". *American Journal of Preventive Medicine* 20.2, pp. 16–66.

Hu, Sean S., Linda Neff, Israel T. Agaku, Shanna Cox, Hannah R. Day, Enver Holder-Hayes, and Brian A. King (2015). *Tobacco Product Use Among Adults — United States, 2013–2014*. Centers for Disease Control and Prevention, pp. 685–691. URL: https://www.cdc.gov/mmwr/volumes/65/wr/mm6527a1.htm (visited on 02/12/2021).

Ioannidis, John P. A. (2005). "Why Most Published Research Findings Are False". *PLOS Medicine* 2.8, e124.

Jones, Andrew M., Audrey Laporte, Nigel Rice, and Eugenio Zucchelli (2015). "Do Public Smoking Bans Have an Impact on Active Smoking? Evidence from the UK". *Health Economics* 24.2, pp. 175–192.

Joossens, Luk, Ariadna Feliu, and Esteve Fernandez (2020). *The Tobacco Control Scale 2019 in Europe*. Association of European Cancer Leagues, Catalan Institute of Oncology. URL: https://www.tobaccocontrolscale.org/TCS2019.pdf.

Kaul, Ashok, Stefan Klößner, Gregor Pfeifer, and Manuel Schieler (2015). *Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together With Covariates*. MPRA Paper. University Library of Munich, Germany. URL: https://EconPapers.repec.org/RePEc:pra:mprapa:83790.

Kennedy, Ryan D. (2010). "Evaluation of the City of Woodstock's Outdoor Smoking By-Law: A Longitudinal Study of Smokers and Non-Smokers". PhD thesis. University of Waterloo.

Ko, Hansoo (2020). "The Effect of Outdoor Smoking Ban: Evidence from Korea". *Health Economics* 29.3, pp. 278–293.

Kueh, Mien T. and Chuan Y. Lin (2020). "The 2018 Summer Heatwaves over Northwestern Europe and Its Extended-Range Prediction". *Scientific Reports* 10.1, pp. 1–18.

Levy, David T., Frank J. Chaloupka, and Joseph Gitchell (2004). "The Effects of Tobacco Control Policies on Smoking Rates: A Tobacco Control Scorecard". *Journal of Public Health Management and Practice* 10.4, pp. 338–353.

Nielsen Corporation (2021). *Retail Measurement*. URL: https://www.nielsen.com/us/en/solutions/measurement/retail-measurement (visited on 02/23/2021).

Niia, Ella, Göran Boëthius, Marie-Louise Luther, Lena Sjöberg, Ingrid Talu, and Pärnilla Zigher (June 1, 2009). "Gör uteserveringarna rökfria". *Svenska Dagbladet*. URL: https://www.svd.se/gor-uteserveringarna-rokfria (visited on 02/11/2021).

Nyberg, Hanna (July 28, 2019). "Lättare än väntat att kontrollera rökarna". *SVT Nyheter*. URL: https://www.svt.se/nyheter/lokalt/vast/lattare-att-kontrollera-rokarna-an-vantat (visited on 02/11/2021).

Odermatt, Reto and Alois Stutzer (2015). "Smoking Bans, Cigarette Prices and Life Satisfaction". *Journal of Health Economics* 44, pp. 176–194.

Organisation for Economic Co-operation and Development (2021a). *Historical Population*. Electronic dataset. URL: https://stats.oecd.org/Index.aspx?DataSetCode=HISTPOP#N (visited on 04/10/2021).

Organisation for Economic Co-operation and Development (2021b). *Non-Medical Determinants of Health: Alcohol Consumption*. Electronic dataset. URL: https://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT# (visited on 04/10/2021).

Organisation for Economic Co-operation and Development (2021c). *Quarterly National Accounts: GDP per capita*. Electronic dataset. URL: https://stats.oecd.org/index.aspx?queryid=66948 (visited on 04/10/2021).

Prop. 2017/18:156. *Ny lag om tobak och liknande produkter*. URL: https://www.regeringen.se/493c3e/contentassets/507611a60846447e817f4fde688be741/ny-lag-om-tobak-och-liknande-produkter-prop-2017_18-156.pdf.

Public Health Agency of Sweden (2014). *Utredning om framtida rökfria miljöer på allmänna platser – i synnerhet där barn vistas*. Public Health Agency of Sweden.

Public Health Agency of Sweden (2019). *New Rules on Smoke-Free Outdoor Environments from 1 July 2019*. URL: https://www.folkhalsomyndigheten.se/publicerat-material/publikationsarkiv/n/new-rules-on-smoke-free-outdoor-environments-from-1-july-2019/ (visited on 02/10/2021).

Ramström, Lars M. and Jonathan Foulds (2006). "Role of Snus in Initiation and Cessation of Tobacco Smoking in Sweden". *Tobacco control* 15.3, pp. 210–214.

Ritzén, Susan (July 29, 2019). "Efter rökförbudet: Bolmandet på uteserveringarna kraftigt reducerat". *SVT Nyheter*. URL: https://www.svt.se/nyheter/inrikes/efter-rokforbudet-svart-att-fa-tag-i-olovliga-rokare (visited on 02/11/2021).

Rubin, Donald B. (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies". *Journal of Educational Psychology* 66.5, pp. 688–701.

SFS 1993:581. *Tobakslag*. URL: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/tobakslag-1993581_sfs-1993-581.

SFS 2010:1622. *Alkohollag*. URL: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/alkohollag-20101622_sfs-2010-1622.

SFS 2017:425. *Lag om elektroniska cigaretter och påfyllningsbehållare*. URL: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-2017425-om-elektroniska-cigaretter-och_sfs-2017-425.

SFS 2018:2088. *Lag om tobak och liknande produkter*. URL: http://rkrattsbaser.gov.se/sfst?bet=2018:2088.

Shiffman, Saul, Michael S. Dunbar, Sarah M. Scholl, and Hilary A. Tindle (2012). "Smoking Motives of Daily and Non-Daily Smokers: A Profile Analysis". *Drug and Alcohol Dependence* 126.3, pp. 362–368.

Socialstyrelsen (2014). *Registeruppgifter om tobaksrökningens skadeverkningar*. URL: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/statistik/2014-3-4.pdf.

State Government of Victoria (2015). *Outdoor Dining - Smoke-Free*. URL: https://www2.health.vic.gov.au/public-health/tobacco-reform/smoke-free-areas/outdoor-dining (visited on 02/25/2021).

Stuber, Jennifer, Sandro Galea, and Bruce G. Link (2008). "Smoking and the Emergence of a Stigmatized Social Status". *Social Science and Medicine* 67.3, pp. 420–430.

The Irish Department of Finance (2019). *General Excise Paper*. URL: http://www.gov.ie/finance/.

Vogel, Martha M., Jakob Zscheischler, Richard Wartenburger, D. Dee, and Sonia I. Seneviratne (2019). "Concurrent 2018 Hot Extremes Across Northern Hemisphere Due to Human-Induced Climate Change". *Earth's Future* 7.7, pp. 692–703.

Vuolo, Mike, Brian C. Kelly, and Joy Kadowaki (2016). "Independent and Interactive Effects of Smoking Bans and Tobacco Taxes on a Cohort of US Young Adults". *American Journal of Public Health* 106.2, pp. 374–380.

Wilson, Lisa M. et al. (2012). "Impact of Tobacco Control Interventions on Smoking Initiation, Cessation, and Prevalence: A Systematic Review". *Journal of Environmental and Public Health* 2012, pp. 1–36.

World Bank (2021a). *Climate Change Knowledge Portal*. Electronic dataset. URL: https://climateknowledgeportal.worldbank.org/download-data (visited on 01/24/2021).

World Bank (2021b). *Population, Total*. Electronic dataset. URL: https://data.worldbank.org/indicator/SP.POP.TOTL (visited on 01/20/2021).

World Health Organisation (2005). "WHO Framework Convention". *WHO Framework Convention on Tobacco Control*. URL: http://www.who.int/tobacco/framework/WHO%7B%5C_%7DFCTC%7B%5C_%7Denglish.pdf.

World Health Organization (2019). *European Tobacco Use: Trends Report 2019*. World Health Organisation. URL: https://www.euro.who.int/__data/assets/pdf_file/0009/402777/Tobacco-Trends-Report-ENG-WEB.pdf (visited on 01/12/2021).

World Health Organization (2020). *Tobacco: Key Facts*. URL: https://www.who.int/news-room/fact-sheets/detail/tobacco (visited on 02/17/2021).

World Health Organization (2021). *The Global Health Observatory*. Electronic dataset. URL: https://www.who.int/data/gho (visited on 02/06/2021).

Zetterqvist, Martina and Mats Ramstedt (2019). *Tobaksvanor i Sverige 2003 – 2019 CAN Rapport 194*. Centralförbundet för alkohol- och narkotikaupplysning, CAN. URL: https://www.can.se/app/uploads/2020/09/can-rapport-194-tobaksvanor-i-sverige-2003-2019.pdf.

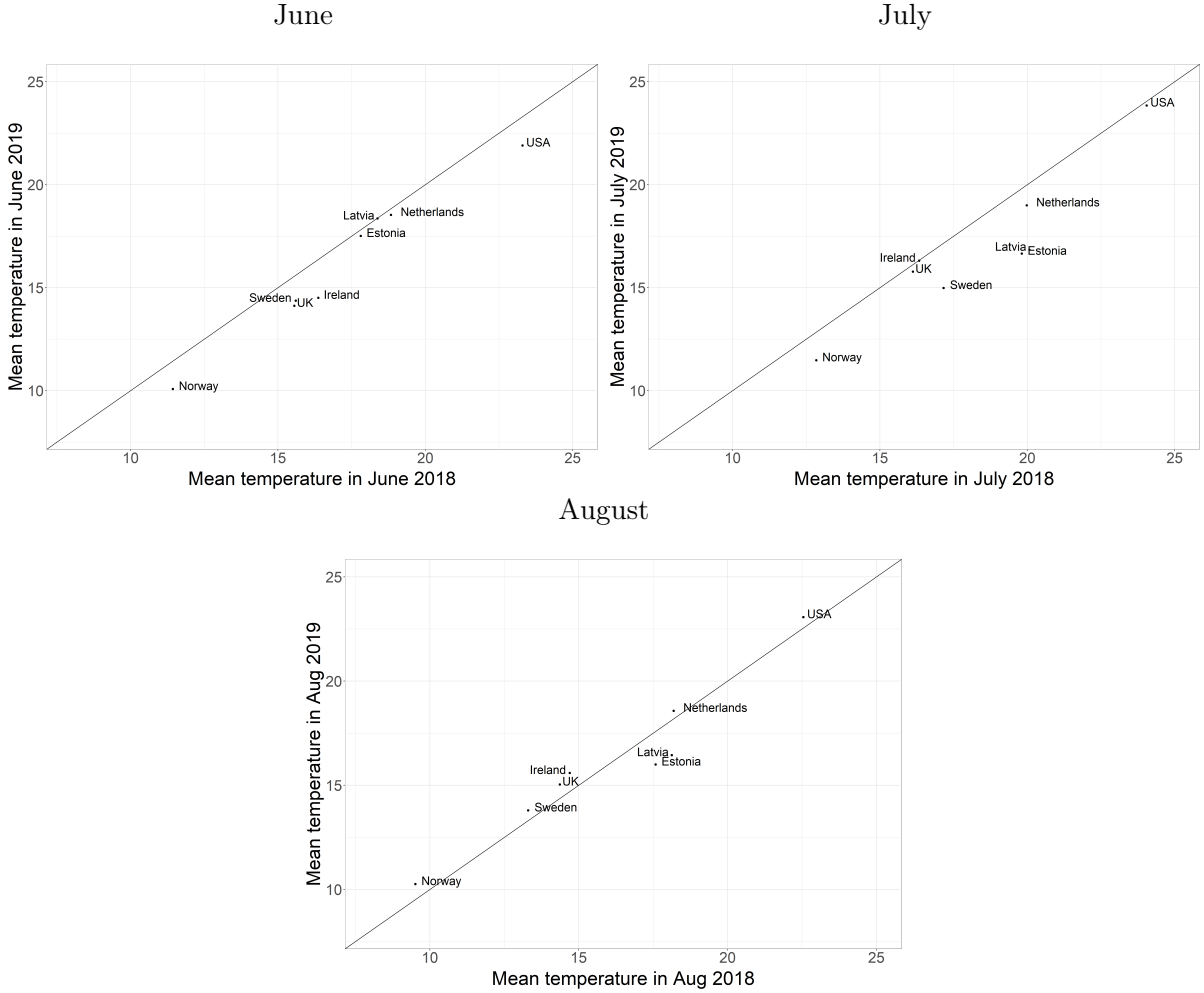
Zhao, Wei, Ningfang Zhou, and Shangfeng Chen (2020). "The Record-Breaking High Temperature over Europe in June of 2019". *Atmosphere* 11.5, pp. 16–20.

Appendices

A Donor pool characteristics

This section contains tables and figures relating to section 5.2 on donor pool characteristics.

Figure A1
Difference in average temperature 2018–19 by summer month



Source: Authors' rendering of data from Berkeley Earth (2021).

Table A1
Overview of major tobacco control policies in the donor pool countries

Estonia	2005: Ban in enclosed public spaces and workplaces. Smoking rooms allowed. Exemption for restaurants and bars 2007: Ban in restaurant and bars
Ireland	2004: Ban in all enclosed public places, workplaces, restaurants and bars 2017: Plain packaging law
Latvia	2005: Ban in workplaces, restaurants and bars, etc. Smoking rooms allowed. 2010: Ban in balconies and windows of multi-story dwelling buildings 2014: Ban in several outdoor areas, such as entrances to enclosed public spaces and parks
Netherlands	2004: Ban in workplaces 2008: Ban in hospitality industry, shopping malls and several other indoor environments. Smoking rooms allowed 2020: Smoking rooms banned
Norway	2004: Ban in enclosed public spaces. Smoking rooms allowed in hospitality industry 2014: Smoking rooms banned
Sweden	2005: Ban in all indoor serving areas. Smoking rooms allowed 2019: Ban in certain outdoor areas
United Kingdom	2007: Ban in enclosed public spaces and workplaces 2017: Plain packaging law
United States	As of 2021, 43 states have bans in workplaces and/or restaurants and bars. Majority introduced in late 1990's and early 2000's.

Note: This overview is only intended to give a brief summary of the countries' progress within tobacco control policies. As such, it leaves out much of the nuances of the laws and does not discuss laws with relatively small impacts.

B Motivation for transforming the outcome variable

In this section, we discuss in more depth our motivations for transforming the outcome variable in this study, cigarette sales per capita, to percentage deviations from the pre-treatment mean. This argument is structured in three parts, motivating why it is appropriate to center the data at zero, and a subsequent fourth part detailing why we go further than simply demeaning the data.

First, due to Nielsen's varying market coverage rate, the structural link that is expected to exist between our outcome series measured in levels and its appropriate predictors is weakened. In particular, the varying market coverage of Nielsen can be seen as random noise or measurement error, but of a severe kind since it differs between orders of ten to twenty percentage points. In a cross-sectional setting, with a large enough sample, the amount of signal to noise might enable us to estimate the link between the predictors and the outcome despite the noise. However, in our setting where the number of donor pool countries is low, this is unlikely to be true. Hence, the predictors will likely end up very uninformative if we match the outcome series on levels, which is indeed confirmed in Section 6 where we perform this kind of analysis as a robustness check; the predictors intended to capture structural determinants of the level get very small predictor weights.

Second, matching synthetic Sweden to the level of the Swedish series is not equivalent to matching it to Sweden's true level since Nielsen measures only some fraction of the market, which undermines the usefulness of doing so even if it would be deemed appropriate for some theoretical reason. While this is not an argument for centering the series at zero per se, it is rather to be considered a constraint imposed on the analysis by the nature of the data. However, we believe that this constraint is not binding, which brings us to our next point.

Third, based on our conceptual understanding of cigarette consumption, we see no reason to expect the response of aggregate cigarette consumption to shocks to depend on the country's level of cigarette consumption. Neither do we have any other reason to expect two countries that are matched on the seasonal pattern and trend in cigarette consumption to develop differently in the seven post-treatment periods that we study. As discussed in Abadie (2020), this is in contrast to studies of e.g. GDP per capita where the convergence theory suggests that the level of countries' series may be important for their expected future development. Therefore, we have no theoretical reason to want to match synthetic and actual Sweden on levels. The conclusion from this discussion is that performing the analysis on the untransformed data would limit the analysis, by restricting the usefulness of predictors and preventing countries with different levels of cigarette consumption but similar seasonal patterns and/or trends from being important parts of synthetic Sweden, while there are no clear benefits.

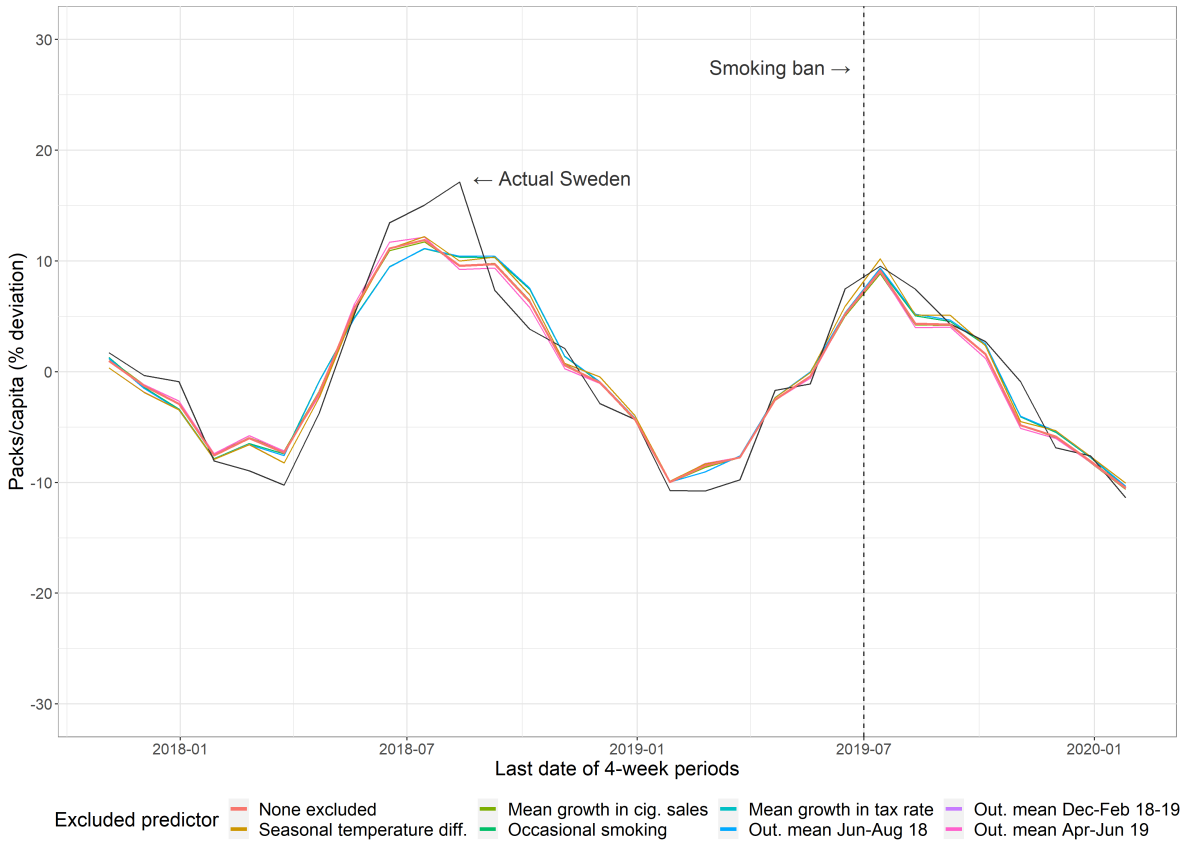
Further, we prefer transforming the series into percentage deviations from the mean over simply demeaning the data, since countries with higher (lower) level of the outcome series mechanically get larger (smaller) seasonal fluctuations. This is likely due to

absolute values of the fluctuations depending on the baseline level of consumption in a given country. Since we have already argued for the actual level of consumption being unimportant in our setting, we see percentage deviations from the mean as a more reasonable measure to use in the analysis. As such, we go one step further than the procedure recommended by Abadie (2020) of demeaning the outcome series when matching on levels is undesirable.

C Synthetic control method

This section contains figures and tables relating to the SC estimator employed in section 5.3.

Figure A2
Sensitivity to exclusion of predictors (default SC)



Note: The black, solid line shows the actual outcomes for Sweden while the colored lines are outcomes estimated by the default SC method, excluding a specific predictor but including the rest.

Figure A3
Country placebo tests (relaxed SC)

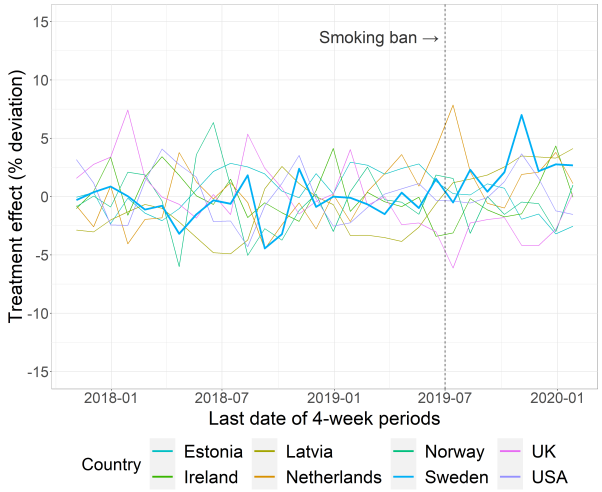


Figure A4
In-time placebo test (relaxed SC)

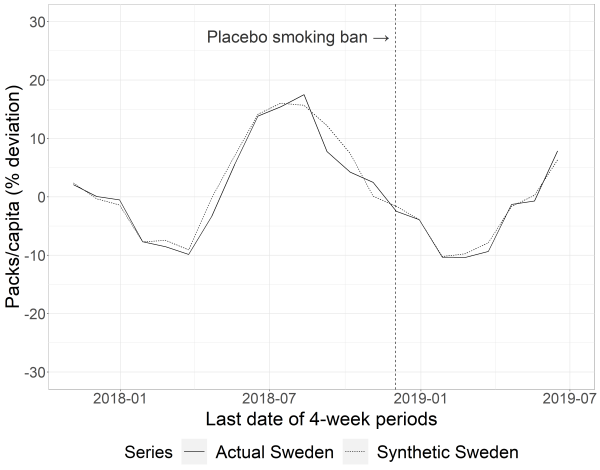


Figure A5
In-time country placebo tests (relaxed SC)

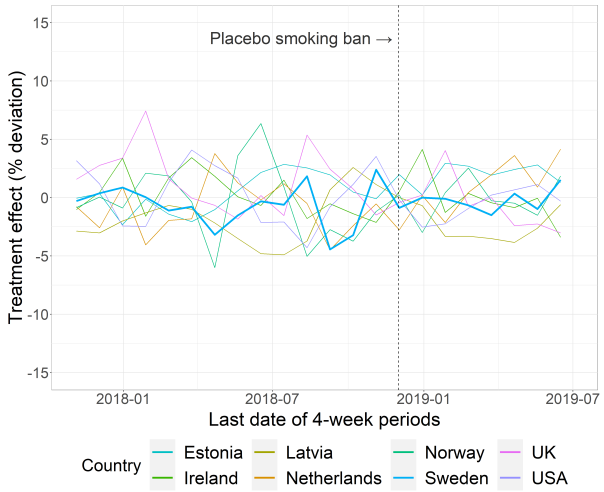
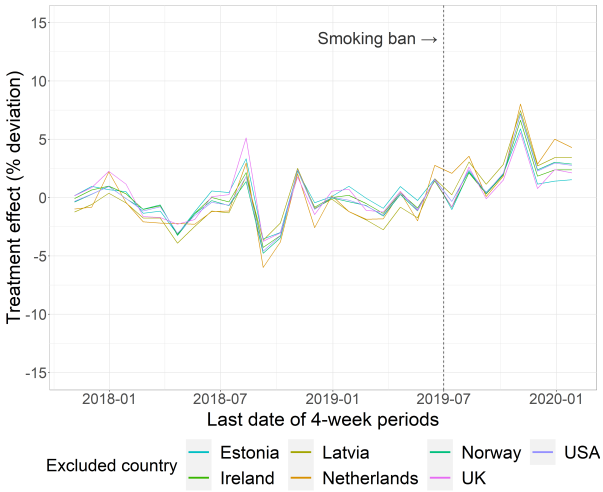


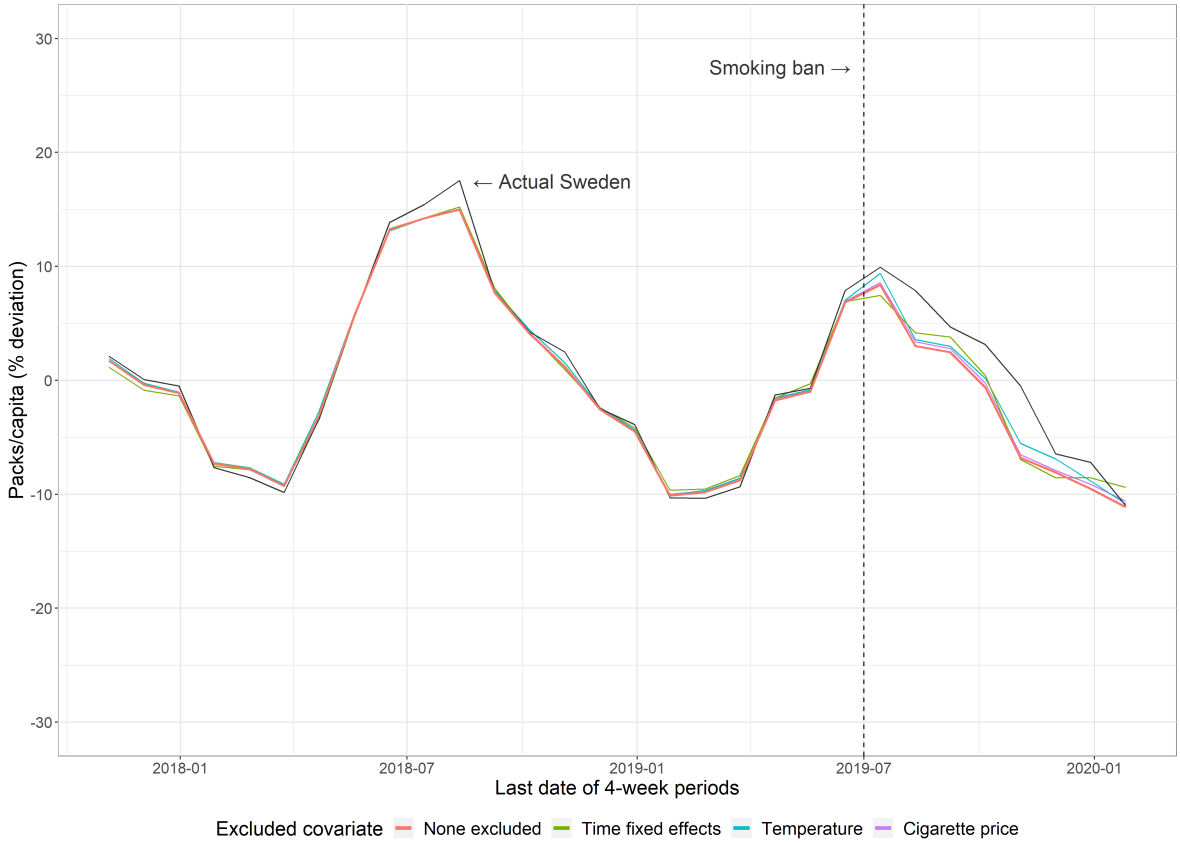
Figure A6
Leave-one-out analysis (relaxed SC)



D Matrix completion with nuclear norm minimization method

This section contains figures relating to the MC estimator presented in 5.4.

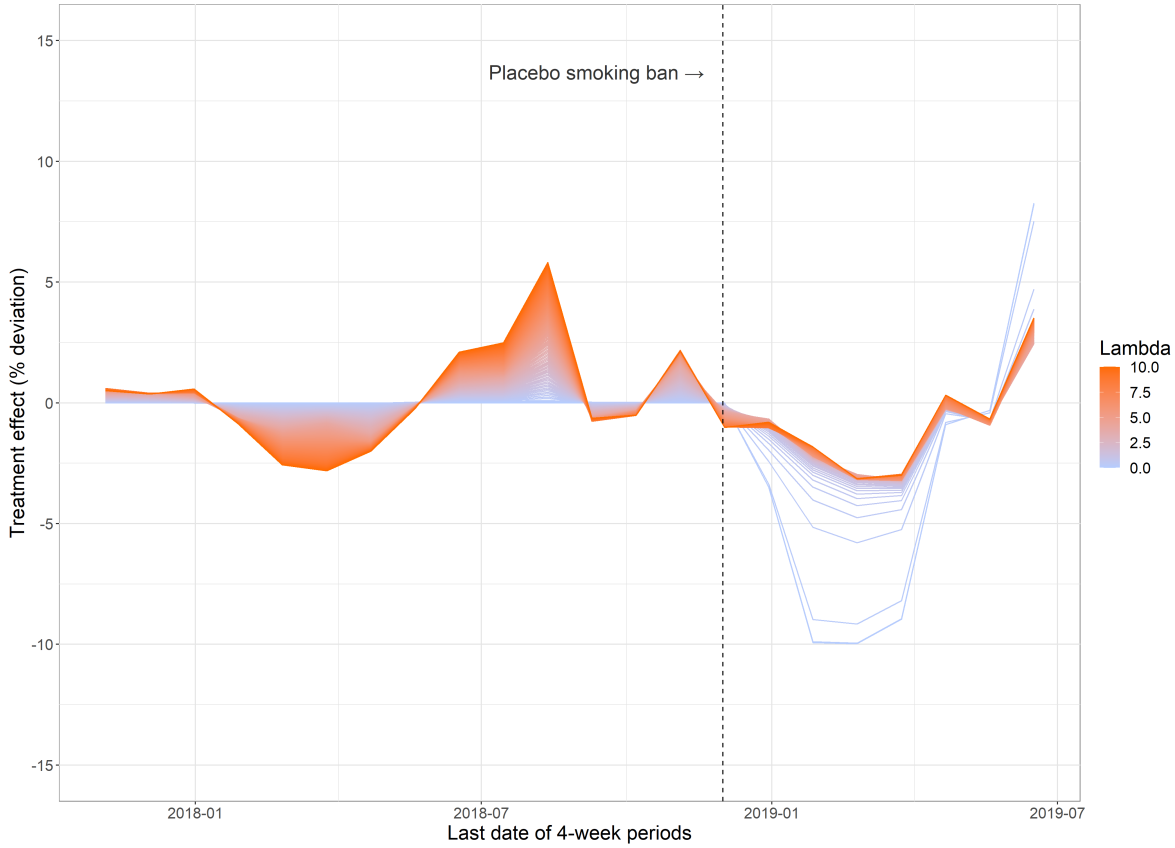
Figure A7
Sensitivity to exclusion of covariates (MC)



Note: The black, solid line shows the actual outcomes for Sweden while the colored lines are outcomes estimated by the MC method, excluding a specific covariate but including the rest.

Figure A8

Sensitivity of in-time placebos to λ, no covariates/fixed effects



Note: The figure shows the treatment gap for Sweden when setting a placebo treatment date to 1 Jan 2019 and letting the λ parameter vary from 0 (blue) to 10 (orange) by steps of 0.1, without including any covariates or fixed effects.