Accident Analysis and Prevention 132 (2019) 105282

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Estimating the effects of a studded footwear subsidy program on pedestrian falls among older adults in Gothenburg, Sweden

Carl Bonander[a,b,*], Robin Holmberg[a]

[a] Centre for Public Safety, Karlstad University, Sweden
[b] Health Metrics Unit, The Sahlgrenska Academy, University of Gothenburg, Sweden

ARTICLE INFO

Keywords:
Vulnerable road users
Falls prevention
Elderly
Road safety
Evaluation
Quasi-experimental

ABSTRACT

We study the effects of a studded footwear subsidy program in Gothenburg, Sweden, where a free pair of anti-slip devices was distributed to all residents aged over 65 years as a pedestrian falls prevention measure. Using a difference-in-differences approach with internal age-based controls, we find evidence of a short-term effect on emergency department visits due to slips on snow and ice during the first year of the intervention (−45% [95% CI: −54, −9] in 2013), which equates to 21.8 injuries prevented (95% CI: 3.34, 39.4). A cost-benefit analysis based on this result suggests that the short-term benefits outweigh the total costs of the intervention (benefit-cost ratio: 6.9 [95% CI: 1.05–12.46]), indicating that this type of subsidy program may be an important tool for the prevention of pedestrian falls among older adults during icy weather conditions. However, replication at other sites is recommended before drawing any strong and general conclusions.

1. Introduction

Falls are the leading cause of unintentional injuries among older adults (Haagsma et al., 2016). Pedestrian falls are also one of the largest transport-related causes of emergency department visits in the Nordic countries, and a majority of these falls are caused by icy road conditions (Elvik and Bjørnskau, 2019; Gyllencreutz et al., 2015; Öberg, 2011; Schepers et al., 2017). To combat the seasonal rise in outdoor falls, the Swedish municipality of Gothenburg introduced a subsidy program that allowed each resident over the age of 65 years to collect a free pair of anti-slip devices ("studded footwear", or ice cleats). Several other Swedish municipalities have adopted this type of program following the implementation in Gothenburg, indicating that it is perceived as an effective intervention.

The evidence regarding the efficacy of anti-slip devices is promising (Berggård and Johansson, 2010; Gao et al., 2008; Gard and Berggård, 2006; Gard and Lundborg, 2001; McKiernan, 2005), but the effects of this type of subsidy program have not yet been documented (Schepers et al., 2017). In this paper, we address this gap using the program in Gothenburg as a case study. We present quasi-experimental effect estimates and conduct an economic evaluation based on our estimates. We also address methodological challenges in evaluating local interventions in which the outcome is affected strongly by local variations in weather conditions.

2. Materials and methods

2.1. Description of the intervention

The subsidy program was performed by the municipality of Gothenburg, Sweden. Starting in 2013, the municipality began to distribute personalized coupons to all residents aged 65 years and above before each winter. The coupons could be exchanged for a free pair of anti-slip devices that can be attached to regular footwear (covering either the entire sole of the shoe, or the heels). The devices could be collected, in person or by proxy, at one of ˜50 local stores. Approximately 82,000 coupons were distributed in the first year (2013). An additional ˜6000 coupons were sent to residents who turned 65 years old in each subsequent year. Each resident was eligible for only one pair of devices throughout the studied period (2013–2016). As of 2016, ˜100,000 coupons had been distributed. Sixty-two percent of the eligible residents cashed in their coupons. The amount of each type of anti-slip device (whole foot or heel device) that was collected is unknown.

2.2. Data

Our outcome measure was the incidence of emergency department visits due to pedestrian falls caused by slipping on snow or ice per

* Corresponding author at: Health Metrics Unit, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, SE-405 30 Gothenburg, Sweden.
E-mail address: carl.bonander@gu.se (C. Bonander).

https://doi.org/10.1016/j.aap.2019.105282
Received 21 January 2019; Received in revised form 17 August 2019; Accepted 21 August 2019
Available online 17 September 2019

100,000 person-years. We collected aggregate emergency department data from the Swedish Traffic Accident Data Acquisition (STRADA) register (held by the Swedish Transport Agency), stratified by year and one-year age group (from 2001 to 2016). We also extracted population data for the same period and groups from Statistics Sweden (the Total Population Register (Ludvigsson et al., 2016)), to calculate person-years as a denominator the for the incidence rates.

We also collected data on the costs of the program, including procurement, distribution and printing costs to perform a cost-benefit analysis (E. Lagerstedt, personal communication, February 3, 2017). The total costs for the program throughout the study period was 1.1 million USD (in 2018), which corresponds to an estimated cost of 13.5 USD per distributed coupon. Following the official recommendations from the Swedish Transport Administration, we set the societal cost of an average (non-fatal) emergency department-treated pedestrian fall injury to 0.35 million USD (The Swedish Transport Administration, 2018). Their estimate is based on losses of quality-adjusted life years (QALYs) from injuries sustained in an average pedestrian fall (based on data from STRADA, the National Patient Register and patient follow-up surveys), which has been converted to monetary values by linking the QALY estimates to the current value of a statistical life (VSL) (Olofsson, Gralén et al. 2016). The VSL was estimated using willingness-to-pay surveys (Olofsson, Persson et al. 2016). Material costs (e.g., healthcare resources and productivity losses) are also added to the estimate (The Swedish Transport Administration, 2018).

The study approved by the Regional Ethics Committee in Uppsala, Sweden (dnr 2018/480).

2.3. Analytical framework

We used a difference-in-differences (DD) approach to estimate the effects of the intervention. The method controls for unobserved, time-invariant confounding and time trends under the assumption that the intervention and control group would have followed the same trend in absence of the intervention (Angrist and Pischke, 2008). The effect estimator can be expressed as:

$$\hat{\delta} = (\bar{y}_{12} - \bar{y}_{11}) - (\bar{y}_{22} - \bar{y}_{21}) \tag{1}$$

where the first term $(\bar{y}_{12} - \bar{y}_{11})$ gives the difference in incidence rate y between the pre- and post-intervention periods in the treatment group, and the second term $(\bar{y}_{22} - \bar{y}_{21})$ gives the same difference in the control group. The causal estimate $\hat{\delta}$ is given by the difference between these two terms (Angrist and Pischke, 2008), and is analogous to the average treatment effect on the treated (ATT). We also estimated relative and cumulative effects (number of injuries prevented). We investigated the average effects for the entire post-period as well as time-varying effects for each post-intervention year to study the longevity of the program effect (see Appendix A for details; R codes and replication files the analyses can also be found on Mendeley (Bonander and Holmberg, 2019).

2.3.1. Control selection

We used internal age-based controls to account for local weather effects and other unobserved determinants of pedestrian safety. Data-driven matching or weighting algorithms that match on pre-intervention trends are often recommended for control selection in DD analyses (Abadie et al., 2010; Bonander, 2018a; O'Neill et al., 2016). We therefore used a data-driven approach to select the most appropriate control ages.

Specifically, we selected the optimal control ages by matching on pre-intervention trends on the outcome using a *k*-nearest neighbours algorithm (Bonander, 2018b). The donor pool consisted of all one-year age groups below 65 years available in the data (range: 12–64 years). To avoid overfitting, we used a leave-one-out cross-validation procedure to determine how many ages to include in the control group (Picard and Cook, 1984). The algorithm is detailed, alongside a theoretical justification for our approach, in Appendix B.

The result from the cross-validation procedure implies that using a combination of six control ages minimizes the root mean squared prediction errors (RMSE) in the pre-intervention data. The resulting matched control group consisted of the ages 47, 50, 52, 54, 62 and 63 years. The pre-intervention RMSE in the matched sample is 0.51 times that of using all untreated ages (Appendix Figure B2).

2.3.2. Uncertainty and sensitivity analyses

We quantified the uncertainty in all estimated parameters using a blocked bootstrap approach (Sills et al., 2015). To avoid changing the estimand by subsetting the treated ages within each bootstrap iteration (as discussed in Appendix B), we treated the intervention ages as a single unit and resampled only the control ages. We then estimated 95% confidence intervals using the percentile method after repeating the entire matching and estimation procedure on 2000 bootstrap resamples.

We also performed placebo studies on untreated ages to estimate the probability of finding an equal or greater effect in age groups where none should be, an inferential method that is often used for synthetic control estimators (Abadie et al., 2010). We repeated this procedure on 2000 random subsamples of the data. In each placebo study, we varied the subsample size *s* from at least 3 to 53 (the full donor pool), and randomly assigned between 1 and *s-2* units to placebo treatment (by coding them as if they were treated), leaving at least two potential controls. We also calculated the ratio between the RMSE in the post-intervention period to the RMSE in the pre-intervention period within each placebo study, which gives an estimate of the effect size relative to how well the estimated counterfactual fits the treated (placebo) unit in the pre-period. Hence, greater weight is assigned to large effects from studies with good pre-intervention fit, and lesser weight to equally large effects from studies with poor fit. We used the same procedure to calculate ratios for each post-intervention time point for the time-varying effect estimates. Based on the empirical distribution of post/pre-RMSE ratios, we then estimated placebo-based "p-values" by calculating the proportion of placebo studies with ratios greater than or equal to the ratio in the actual treatment group (65+ years) (Abadie et al., 2010). The "p-values" can be interpreted as the probability of finding an effect estimate elsewhere in the data that is equal to, or greater than, the effect found in the main analysis, given an equal pre-intervention fit.

To check that no single control age contributed extremely to the results, we also performed an extensive leave-*k*-out sensitivity analysis in which we re-ran the analysis *k* times, iteratively leaving out the best control ages until the pre-intervention RMSE was 1.5 times worse than in the original model (Abadie et al., 2015; Bonander, 2018a). We repeated this procedure on smaller pre-intervention time spans to also check for robustness against the selection of the pre-intervention period. We note that the stopping rule is arbitrarily chosen. It does, however, have some appeal as it is ˜0.75 times the prediction error from the unmatched analysis, i.e., a fit that is half as good as the optimal control unit relative to using all available control ages.

2.3.3. Cost-benefit analysis

We calculated the benefit-cost ratio (BCR) as our primary measure for the economic evaluation, and used the bootstrapped uncertainty estimates from the effects (Section 2.3.2) to perform probabilistic sensitivity analyses and quantify confidence intervals for the economic measures (Briggs et al., 2012). Specifically, we calculated the BCR, for the primary analysis and selected sensitivity analyses, as follows:

$$BCR = \frac{\hat{\alpha}_t * b}{c}, \tag{2}$$

where $\hat{\alpha}_t$ is the estimated number of injuries prevented in period t; b is the cost per pedestrian fall injury and c is the total cost for the program (thus, a BCR $>$ 1 implies that the intervention is cost-beneficial). We also quantified a break-even effect to evaluate how many injuries need to be prevented in order for the intervention to be cost-beneficial given the assumed cost per injury. This quantity is given by c/b, as can be derived by setting BCR to 1 solving for $\hat{\alpha}_t$ in Eq. (2). In addition to calculating the break-even effect for the intervention in Gothenburg (where ˜82,000 coupons were distributed), we also quantified a break-even effect in injuries prevented per 10,000 coupons to provide a more generalizable measure.

3. Results

A total of 2679 emergency department visits due to pedestrian falls caused by slipping on snow or ice were reported to the STRADA register by hospitals in Gothenburg during the study period (2001–2016). Twenty-eight percent (28.5%, n = 763) of the injured individuals were above the treatment age threshold (65+ years), and the mean age was 54.4 years (range: 12–97). Overall, the recorded observation time was 7,050,825 person-years, indicating an injury rate of 38 per 100.000 person-years in the general population of Gothenburg. The rate in the treated age group was roughly twice as large as the rate in the full untreated age range (62.53 vs. 32.86 per 100.000 person-years).

3.1. Intervention effects

The time-varying results from the DD analysis are presented in Fig. 1. The observed injury rates and the estimated counterfactual based on the matched controls are the presented in panel (a). The effect estimates, including pre-intervention periods for reference, are presented in panel (b). Panel (a) also includes a counterfactual based on the full untreated age range for comparison. Visual inspection of the data in the pre-intervention period confirms that the full untreated age range does not follow the treatment group during warmer and colder years (2010 was an especially cold year). The matched controls, on the other hand, follow the treatment group closely throughout the pre-intervention period, indicating that they are a more appropriate control group.

Focusing on the post-period, visual inspection of Fig. 1 does not provide any convincing evidence of a long-term effect of the intervention, and the average effect estimates are indeed negative but insignificant at the 5% level (Table 1). However, there is a large reduction in injury rates in the intervention group during the first post-intervention year (2013), which might imply a short-term effect (Fig. 1 & Table 1). The estimates indicate that 21.8 (95% CI: 3.34, 39.4) injuries were prevented as a result (−45% [95% CI: −54, −9] in relative terms). This suggests that the effect per 10,000 distributed coupons is −2.65 pedestrian fall injuries (95% CI: −4.79, −0.40), and that 3769 (95% CI: 2086, 24,702) coupons need to be distributed to prevent one injury.

Table 1
Average and time-varying estimates for the effect of the studded footwear subsidy program in Gothenburg, Sweden on emergency department visits due to pedestrian falls on snow or ice.

Estimator	Effect estimate	95% CI	Placebo p-value
Average post-effect (RD)	−6.54	−22.2, 4.77	0.12
Cumulative effect	21.7	−15.6, 74.4	–
Relative effect (RR)	0.90	0.73, 1.08	–
Year-by-year (RD)			
2013	−34.4	−48.8, −4.18	0.024
2014	9.00	−13.6, 19.7	0.37
2015	11.8	−5.44, 22.94	0.24
2016	−13.3	−49.8, 8.89	0.39

Notes: Theoutcome is emergency department visits after pedestrian falls due to slipping on snow or ice. Rate differences (RD) are expressed rates per 100.000 person−years. Averge post-period effects are also expressed as the cumulative number of injuries prevented ("cumulative effect") and in relative terms (rate ratio, RR). Confidence intervals (CIs) were obtained using a blocked bootstrap approach, re-estimating the entire procedure on 2000 bootstrap resamples of all potential control ages (the donor pool)· Placebo p-values were obtained by analysing randomly assigned, fake interventions in 2000 random subsamples of the donor pool.

3.2. Economic analysis

The estimate for the first post-intervention year alone implies that the benefits outweigh the (total program) costs by a ratio of 6.9 (95% CI: 1.05–12.46), and that the intervention is cost-beneficial in 97.6% of the 2000 bootstrap iterations used to quantify uncertainty in the effect estimates. The break-even analysis indicates that 3.16 injuries would have to be prevented for the benefits to match the costs of the intervention in Gothenburg, assuming the cost of injury is correctly estimated. Expressed in more general terms, break-even effect is -0.38 pedestrian falls per 10,000 coupons (assuming the cost per coupon is the same as in Gothenburg). An editable Excel file, which can be used to change inputs for contexts where program costs or the cost of injury differs, can be found at Mendeley (Bonander and Holmberg, 2019).

3.3. Sensitivity analyses

The results from the sensitivity analyses consistently indicate an effect in 2013, even though the pre-intervention fit becomes progressively worse as more pre-intervention years are dropped (Appendix Fig.

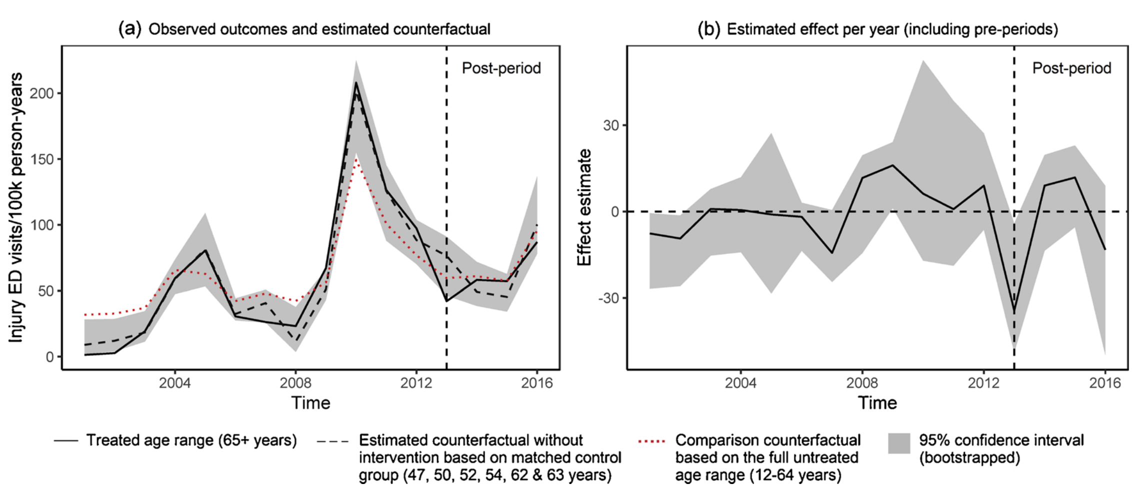

Fig. 1. Observed outcomes, counterfactuals (panel a) and estimated effects of the anti-slip device subsidy program (panel b) on emergency department (ED) visits due to pedestrian falls on snow/ice in Gothenburg, Sweden. The vertical reference line shows the start of the intervention.

C1). The median estimate for 2013 from all sensitivity analyses is 21.96 injuries prevented (min: 12.03, max: 46.09), which is close to the main result. The BCR based on the smallest effect estimate is 3.8.

As the intervention was implemented before the winter season at the end of 2013, a potential issue with the current data, which is stratified by year, is that the drop may have occurred in the beginning of the year. We therefore re-ran the results using winters (e.g., Oct 2013 to March 2014) as time periods. Unfortunately, the matching algorithm performed poorly on this data (the pre-intervention RMSE was two times worse than in the original analysis). The estimate still indicated a drop in injuries, but was no longer statistically significant (11.7 injuries prevented [95% CI: −17.6, 27.7] in the winter of 2013/2014). Using this data, the intervention was only cost-beneficial in 70% of the bootstrap iterations. However, manually applying the same matched control as in the main analysis yielded a similar point estimate to the main results (20.7 injuries prevented).

4. Discussion

Our results imply that the subsidy program was successful in reducing the incidence of pedestrian fall injuries, which is in line with previous research on the effects of anti-slip devices (Berggård and Johansson, 2010; Gard and Berggård, 2006; McKiernan, 2005). The effects, however, appear short-lived, which is consistent with data on the longevity of the effect of information campaigns (Finseraas et al., 2017; Gerber et al., 2011), but may also indicate that the distributed devices had a short life span (e.g., due to poor quality or incorrect use). Negative experiences (e.g., if they are hard to take on/off and difficult to carry) can also potentially influence the uptake and continued use of anti-slip devices (Gard and Lundborg, 2001; Gard and Berggård, 2006). Thus, potential strategies to increase the longevity of similar interventions could be to remind and/or re-supply the population after one year and improve the quality of the distributed anti-slip devices to match the needs of the target population. Unfortunately, our data does not give sufficient insight into which of these, if any, is the most viable option, which makes this an important avenue for future research.

Still, the results imply that the short-term effects were sufficient to motivate the program economically, assuming our estimates are causal. Given that our study is quasi-experimental, and that the sensitivity analyses cast some doubt on the main results, it may be ill-advised to infer causality without further evidence. As a sanity check, we compared our results to McKiernan (2005), who conducted a randomized experiment on 109 older adults in Wisconsin, USA (age range: 65–96 years). The estimated rate ratio for injurious falls was 0.1 (95% CI: 0.02–0.53). According to our data, 64% of the eligible residents collected a pair of anti-slip devices in 2013. This implied relative effect based on the estimate from McKiernan (2005) is then (0.1–1)*0.64 = −58% (95% CI: −63%, −30%), which is larger than our estimate for 2013 (−45%). Hence, our estimate is within a reasonable range. Even so, we cannot rule out biases such as concurrent interventions that disproportionately affect older pedestrians or spillover effects on younger age groups, which motivates replication at other sites.

Other limitations to our study include that the cost of injury is based on QALY-to-VSL conversion and material costs for an average pedestrian fall injury observed in the STRADA database. QALY-to-VSL estimates can be highly variable between studies (Ryén and Svensson, 2015), and the current estimate does not account for the fact that the analysed injuries reflect an older-than-average population. Nonetheless, it is the official estimate currently recommended for decision-making by the Swedish Transport Administration (2018), which makes the estimate relevant for the context. Another limitation is that the data and estimates are generally imprecise. We therefore refrained from conducting any moderation or subgroup analyses to check for heterogeneity in the effects of the program. For instance, with more data we would have liked to investigate gender differences, as previous research has shown that older women are more likely to be injured in pedestrian falls in Sweden (Öberg, 2011), and there is qualitative evidence that suggests that women have more positive attitudes towards using anti-slip devices than men (Pohl et al., 2015). Hence, there is reason to suspect that men and women may react differently to interventions of this kind.

External validity is also a potential concern. Even if our estimates are internally valid, the effects of anti-slip device programs are likely to be heterogeneous depending on climate and population characteristics. Program effects and costs may also vary depending on delivery and dose. Hence, replication is needed to assess the generalizability of our results. Based on data from newspapers and municipal websites, we estimate that roughly 70 municipalities in Sweden have implemented studded footwear subsidy programs after Gothenburg, which may allow for a more comprehensive evaluation in the near future. This could also allow for a more detailed investigation into factors that impact the size of the program effects, and whether certain subgroups (e.g., by gender and age) benefit more than others. It would also be interesting to study the causes of the drop in effectiveness after the first year in greater detail, especially if this pattern emerges in other intervention communities. Conducting a process evaluation, including target population surveys of attitudes and experiences with these programs, could probably generate valuable input for such analyses. Estimating program effects on walking would also be of considerable interest (Berggård and Johansson, 2010; McKiernan, 2005), as this may be an additional health benefit to this type of intervention (Andersen et al., 2000).

5. Conclusions

Implementing anti-slip device subsidy programs may be an efficient way to prevent pedestrian falls among older adults during the winter. Further research is needed to assess the generalizability, and to re-affirm the validity, of our results.

Funding

This work was supported by a research grant from Familjen Kamprads Stiftelse (grant number 20180067).

Declaration of Competing Interest

None.

Acknowledgements

We thank Eva Lagerstedt for providing data on the implementation and costs of the program, and Johanna Gustavsson and Mikael Svensson for providing valuable discussions as input for study.

Appendix A. Estimators

Our goal was to estimate the effect of the program on pedestrian fall injuries. Our main outcome measure was an incidence rate. The standard method for dealing with rates over time is to sum the numerator (n fall injuries) and the denominator (population) over the study period, the latter then becoming a measure of person-time instead of population. In a difference-in-differences (DD) context, we sum these within each group and pre-post period. For notational simplicity, let $F_{it} = \sum f_{it}$ denote the sum (frequency) of injury events in group i and period t, where $i = 1$ is the treated age group and $i = 0$ is the control, and $t = T0$ is the pre-intervention period and $t = T1$ is the post-period. We define exposure in terms of person-time, $E_{it} = \sum e_{it}$, analogously. The estimator for the average treatment effect on the treated (ATT) becomes

$$\hat{\delta}_1 = \left(\frac{F_{1,T1}}{E_{1,T1}} - \frac{F_{1,T0}}{E_{1,T0}}\right) - \left(\frac{F_{0,T1}}{E_{0,T1}} - \frac{F_{0,T0}}{E_{0,T0}}\right). \tag{A1}$$

We also consider time-varying effects. In a slight abuse of notation, we will now define $y_{it} = f_{it}/e_{it}$ as the rate in group i at a single time point and $\bar{y}_{it} = F_{it}/E_{it}$ as the period-wise (pre or post) rate, returning to a more classic DD notation by treating the count data structure as implicit. The time varying effect at time t is then given by

$$\hat{\delta}_{1t} = y_{1,t} - [(\bar{y}_{1,T0} - \bar{y}_{0,T0}) + y_{0,t}], \tag{A2}$$

where $y_{1,t}$ is the rate in the treated unit at time t; $(\bar{y}_{1,T0} - \bar{y}_{0,T0})$ is the difference in rate between the treatment and control group over the entire pre-intervention period, and $y_{0,t}$ is the rate in the control group at time t. The term $[(\bar{y}_{1,T0} - \bar{y}_{0,T0}) + y_{0,t}]$ standardizes the control group to the pre-intervention level of the treated unit, and serves as the estimated, time-varying counterfactual.

We also obtain estimates of the cumulative number of injuries prevented by calculating and summing the term $\hat{\delta}_{1t}e_{1t}$ over the post-period. Let $\hat{CE}$ denote this cumulative effect. From here, we also calculate a relative effect measure using

$$\hat{RR} = \frac{F_{1,T1}}{F_{1,T1} - \hat{CE}}, \tag{A3}$$

where $F_{1,T1}$ is the observed sum of injury events in the post-period in the treatment group, and $F_{1,T1} - \hat{CE}$ is estimated counterfactual number of events without the intervention.

Appendix B. Control selection

Our analysis relied on what we called an "unconstrained" selection of age controls. This appendix expands on this idea and discusses the theoretical implications of defining controls in this manner. To our knowledge, the most common strategies for control selection in age-based difference-in-differences analyses are to either use the entire untreated age group, or some restricted age span closer to the treated age group. However, other options could be explored with a data-driven approach, as exemplified in Fig. B1.

The first row shows the basic choice, which is to use all available ages in the analysis ("the full range option"). The other options involve using some subset of ages. We divide these options into two archetypes that either are (examples 2–5), or are not (examples 6–9), subjected to adjacency constraints. Within each of the two archetypes, we have four different options: to subset the untreated group only, to subset the treated group only, or to subset both at the same time (equally or unequally).

Control selection algorithm (details)

We selected the optimal control ages by matching on pre-intervention trends using a k-nearest neighbours algorithm (Bonander, 2018b). Specifically, the algorithm subtracts the pre-intervention average of the outcome in each one-year age group i to isolate the temporal variation in the data. To measure similarity in trends, it then determines the Euclidean distance from each control age (in this case, each one-year age group) to the

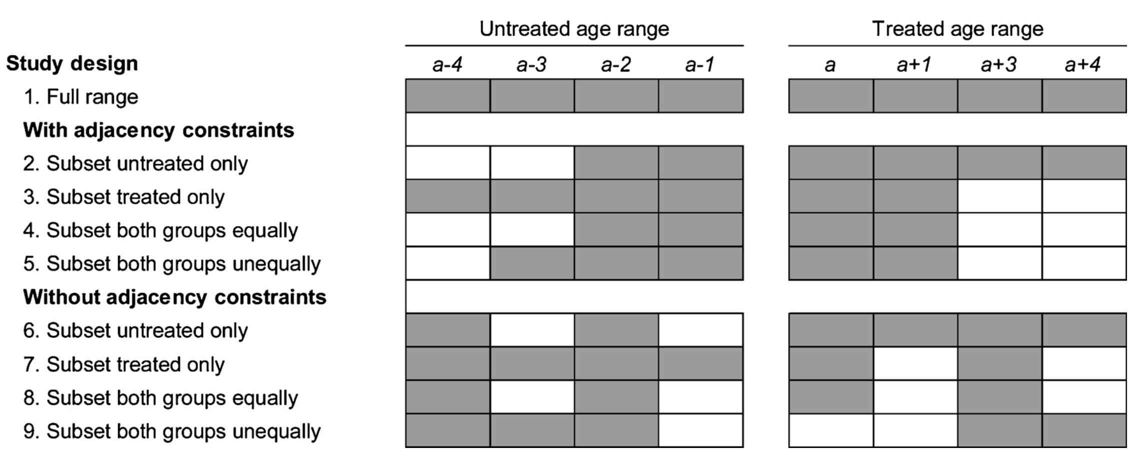

Fig. B1. Examples of different options for constructing age groups in difference-in-differences analyses (shaded cell indicates that the age a is included).

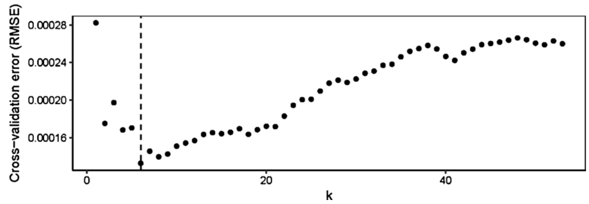

Fig. B2. Root mean squared cross-validation errors (RMSE) for each possible choice of k (the number of controls to include) in the nearest-neighbours matching procedure, based on leave-one-out cross-validation in the pre-intervention period. The maximum (k = 53) reflects the RMSE for all untreated ages (12–64 years). The vertical line indicates the number of nearest neighbours (matched controls) that minimizes the error (k = 6). The final match implies that a control unit consisting of the ages 47, 50, 52, 54, 62 and 63 years is the best match for the intervention ages (65+ years).

trends in treated age range. It then ranks the potential controls from $k = 1$ to $k = K$, where 1 is the untreated age with the most similar pre-trend and K is the least similar. We used a leave-one-out cross-validation procedure to determine how many ages to include in the control group (Picard and Cook, 1984). Specifically, we iteratively dropped each pre-intervention time point from the analysis to serve as an out-of-sample test point, re-running the matching procedure on the remaining time points. We did this for each pre-intervention time point and evaluated all possible choices of k in each iteration. We then selected the number of controls k that minimized the out-of-sample root mean squared prediction errors (RMSE) based on the average of the test points, and used this number to obtain a final match using the full dataset. The RMSE for all possible choices of k are presented in Fig. B2, which shows that $k = 6$ minimizes the out-of-sample prediction error.

Effects of the control selection procedure on the causal estimand

As detailed in the previous section, we employed a data-driven approach to selecting age controls that allows for any unconstrained (irrational) combination of control ages. We here provide proof that this makes no difference for the causal estimand (we are still able to estimate the average treatment effect on the treated, ATT, as if we had used any other, constrained age group as controls). We can write the following general estimand for all options in Fig. B1:

$$E[\delta | D = 1, A^1, t] = E[Y_i^1 - Y_i^0 | D = 1, A^1, t] \tag{B1}$$

where $E[\delta_t | D = 1, A^1, t]$ the conditional ATT at time t; $Y_i^1 - Y_i^0$ is the difference in potential outcomes for each individual i in the population, where Y_i^1 is observed outcome and Y_i^0 is the counterfactual; D is an indicator for the ages eligible for treatment, and A^1 defines the age strata among the treated ages that is included in the analysis (A^0, used below, is defined analogously for controls). If all treated ages are included, we can remove this term. Hence, for any option that includes the full treated age range (examples 1, 2 and 6 in Fig. B1), we can obtain unbiased estimates of $E[\delta | D = 1, t]$, the unconditional ATT defined for some time period t, assuming the identifying assumptions of the DD analysis hold. If we subset the treated group, however, we change the estimand to the conditional ATT for the selected subset (unless the treatment effect is homogenous for all treated ages). This estimand may be of limited policy interest, especially if the adjacency constraints are dropped for the treated (as in examples 7, 8 and 9 in Fig. B1). We therefore focused on subsetting the untreated age group in our study. However, subsetting the treated could be viable option if substantial pre-intervention prediction errors remain even after matching, and one is willing to concede to a more internally valid local estimate at the (likely) loss of generalizability to the entire treatment population. We note that it may be important to keep this in mind when quantifying the uncertainty using bootstrapping as well, and to avoid bootstrapping the treatment ages if the effects are not homogenous within this group (as this will change the estimand within each run of the bootstrap).

An important aspect here is that the estimand does not depend on the selection of untreated ages. The bias of the DD estimator does, however, assuming that some controls are more valid than others. We can write the estimator for a two period case, where $t = 1$ and $t = 0$ define the pre and post periods, as follows:

$$E[\hat{\delta} | D = 1, A^1, t{=}1] = \{E[Y_i | D = 1, A^1, t{=}1] - E[Y_i | D = 1, A^1, t{=}0]\} - \{E[Y_i | D = 0, A^0, t{=}1] - E[Y_i | D = 0, A^0, t{=}0]\}, \tag{B2}$$

where Y_i is the realized (observed) outcome in age group i, and $E[Y_i | D = 0, A^0, t = 1] - E[Y_i | D = 0, A^0, t = 0]$ is used as a substitute for the true counterfactual time difference among the treated $E[Y_i^0 | D = 1, A^1, t = 1] - E[Y_i^0 | D = 1, A^1, t = 0]$ in absence of the intervention, which is unobservable. We can then decompose the bias of the estimator into the difference between substitute and the true counterfactual:

$$\begin{aligned} bias_{\hat{\delta}} = E[\hat{\delta} | D = 1, A^1, t{=}1] - E[Y_i^1 - Y_i^0 | D = 1, A^1, t{=}1] = \{E[Y_i^0 | D = 0, A^0, t{=}1] - E[Y_i^0 | D = 0, A^0, t{=}0]\} \\ - \{E[Y_i^0 | D = 1, A^1, t{=}1] - E[Y_i^0 | D = 1, A^1, t{=}0]\}, \end{aligned} \tag{B3}$$

i.e., the difference in the time difference in Y_i^0 among the controls in age strata A^0 and the counterfactual time difference in Y_i^0 among the treated. Hence, assuming $E[Y_i^0 | D = 0, A^0, t{=}1] - E[Y_i^0 | D = 0, A^0, t{=}0] = E[Y_i^0 | D = 1, A^1, t{=}1] - E[Y_i^0 | D = 1, A^1, t{=}0]$, i.e., common trends in absence of the intervention between the groups defined by A^0 and A^1, we have that $bias_{\hat{\delta}} = 0$ (given that there are no spillover effects on A^0). If there is a combination of treated and untreated ages in the population where this assumption holds, causal identification can be achieved by either conditioning on A^0, A^1 or both, as long as they are correctly selected. This motivates the data-driven control selection approach used in our paper.

Appendix C. Sensitivity analyses

See Fig. C1

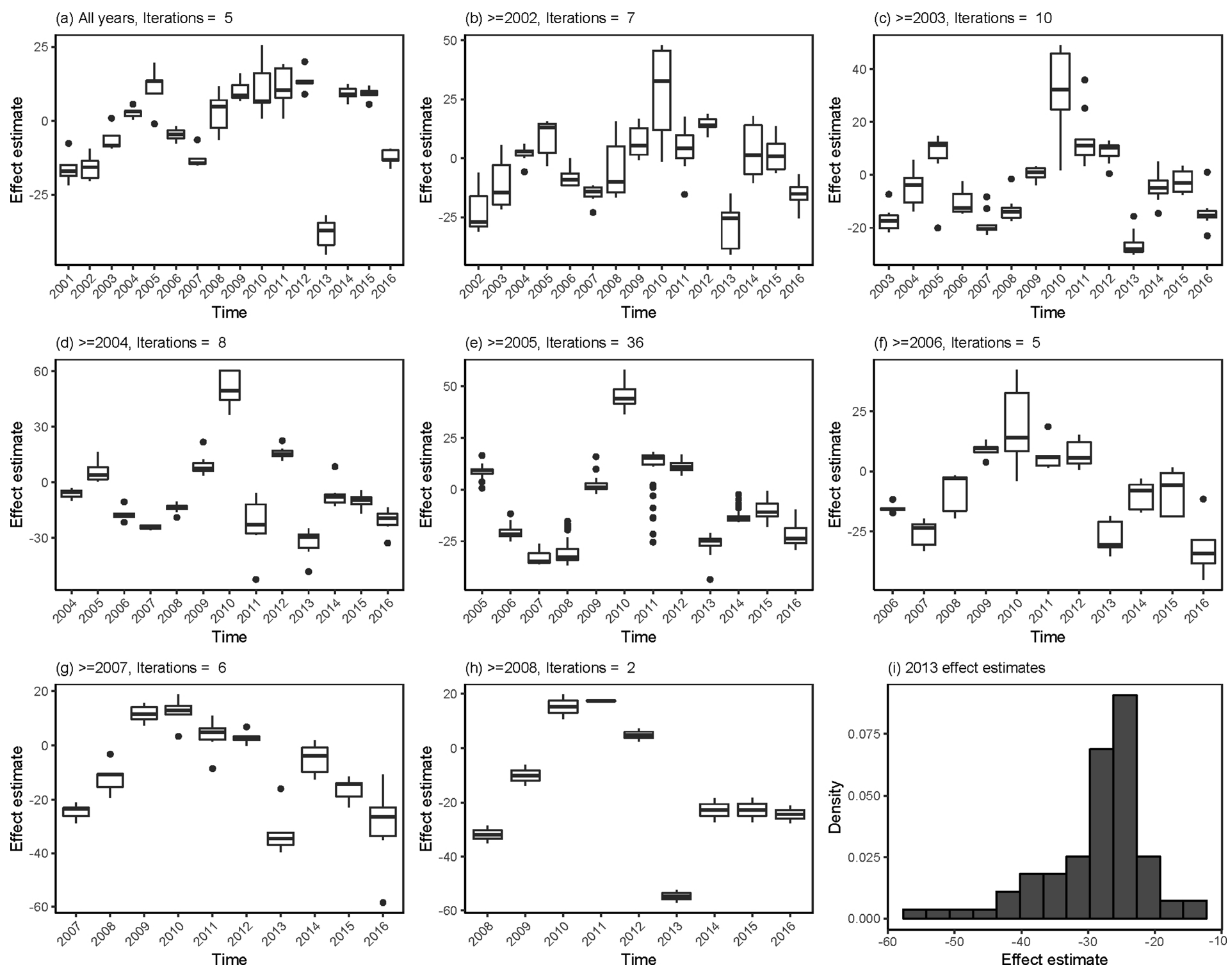

Fig. C1. Box plots showing the distribution of effect estimates from the leave-k-out sensitivity analysis conducted on smaller and smaller pre-intervention time spans (subplots a-h). Within each subplot in the range a-h, we iteratively leave out the best fitting of the matched controls (specifically, with the lowest cross-validated RMSE) until an average RMSE of 75% of the size of an unmatched analysis is reached, in order to probe the sensitivity of the results to the inclusion of influential control units. In each subplot in the range a-h, the time-specific boxes show the median (vertical line), interquartile range (box) and outliers (dots) of the time-specific effect estimates from the sensitivity analysis iterations (including the pre-period for reference. Finally, subplot (i) shows a histogram of the estimated effect in 2013 from all sensitivity analyses. All estimates are presented in incidence rates per 100,000 person-years.

References

Abadie, A., Diamond, A., Hainmueller, J., 2010. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. J. Am. Stat. Assoc. 105 (490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746.

Abadie, A., Diamond, A., Hainmueller, J., 2015. Comparative politics and the synthetic control method. Am. J. Pol. Sci. 59 (2), 495–510. https://doi.org/10.1111/ajps.12116.

Andersen, L.B., Schnohr, P., Schroll, M., Hein, H.O., 2000. All-cause mortality associated with physical activity during leisure time, work, sports, and cycling to work. Arch. Intern. Med. 160 (11), 1621–1628.

Angrist, J.D., Pischke, J.-S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Berggård, G., Johansson, C., 2010. Pedestrians in wintertime—Effects of using anti-slip devices. Accid. Anal. Prev. 42 (4), 1199–1204. https://doi.org/10.1016/j.aap.2010.01.011.

Briggs, A.H., Weinstein, M.C., Fenwick, E.A., Karnon, J., Sculpher, M.J., Paltiel, A.D., 2012. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM modeling good research practices task force working group–6. Med. Decis. Mak. 32 (5), 722–732.

Bonander, C., 2018a. Compared with what? Estimating the effects of injury prevention policies using the synthetic control method. Inj. Prev. 24 (Suppl. 1), i60–i66. https://doi.org/10.1136/injuryprev-2017-042360.

Bonander, C., 2018b. Idd: Data-Driven Incidence Difference-in-Differences Estimators (Version 0.4) [R Package]. Retrieved from. https://github.com/carlbona/idd.

Bonander, C., Holmberg, R., 2019. Data for: Estimating the effects of a studded footwear subsidy program on pedestrian falls among older adults in Gothenburg, Sweden. [dataset]. Mendeley Data v1.

Elvik, R., Bjørnskau, T., 2019. Risk of pedestrian falls in Oslo, Norway: relation to age, gender and walking surface condition. J. Transp. Health 12, 359–370. https://doi.org/10.1016/j.jth.2018.12.006.

Finseraas, H., Jakobsson, N., Svensson, M., 2017. Do knowledge gains from public information campaigns persist over time? Results from a survey experiment on the Norwegian pension reform. J. Pension Econ. Finance 16 (1), 108–117. https://doi.org/10.1017/S1474747215000098.

Gao, C., Holmér, I., Abeysekera, J., 2008. Slips and falls in a cold climate: underfoot surface, footwear design and worker preferences for preventive measures. Appl. Ergon. 39 (3), 385–391. https://doi.org/10.1016/j.apergo.2007.08.001.

Gard, G., Berggård, G., 2006. Assessment of anti-slip devices from healthy individuals in different ages walking on slippery surfaces. Appl. Ergon. 37 (2), 177–186. https://doi.org/10.1016/j.apergo.2005.04.004.

Gard, G., Lundborg, G., 2001. Test of Swedish anti-skid devices on five different slippery surfaces. Accid. Anal. Prev. 33 (1), 1–8.

Gerber, A.S., Gimpel, J.G., Green, D.P., Shaw, D.R., 2011. How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment. Am. Polit. Sci. Rev. 105 (1), 135–150.

Gyllencreutz, L., Björnstig, J., Rolfsman, E., Saveman, B.-I., 2015. Outdoor pedestrian fallrelated injuries among Swedish senior citizens – injuries and preventive strategies. Scand. J. Caring Sci. 29 (2), 225–233. https://doi.org/10.1111/scs.12153.

Haagsma, J.A., Graetz, N., Bolliger, I., Naghavi, M., Higashi, H., Mullany, E.C., et al., 2016. The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study 2013. Inj. Prev. 22 (1), 3–18. https://doi.org/10.1136/injuryprev-2015-041616.

Ludvigsson, J.F., Almqvist, C., Bonamy, A.-K.E., Ljung, R., Michaëlsson, K., Neovius, M., et al., 2016. Registers of the Swedish total population and their use in medical research. Eur. J. Epidemiol. 31 (2), 125–136. https://doi.org/10.1007/s10654-016-0117-y.

McKiernan, F.E., 2005. A simple gait-stabilizing device reduces outdoor falls and non-serious injurious falls in fall-prone older people during the winter. J. Am. Geriatr. Soc. 53 (6), 943–947. https://doi.org/10.1111/j.1532-5415.2005.53302.x.

Öberg, G., 2011. Skadade fotgängare: Fokus på drift och underhåll vid analys av sjukvårdsregistrerade skadade i STRADA [In Swedish] (VTI Report No. 705). Retrieved from: http://www.diva-portal.org/smash/get/diva2:670581/FULLTEXT01.pdf [2019-08-16]. .

Olofsson, S., Gralén, K., Macheridis, K., Welin, K.O., Persson, U., Hultkrantz, L., 2016a. Personskadekostnader och livskvalitetsförlust till följd av vägtrafikolyckor och fotgängarolyckor singel. Fullständig rapport [In Swedish] (IHE Report No. 2016:5).

Olofsson, S., Persson, U., Hultkrantz, L., Gerdtham, U., 2016b. Betalningsviljan för att minska risken för icke-dödliga och dödliga skador i samband med vägtrafikolyckor–en studie med kedje-ansats [In Swedish] (IHE Report No. 2016:7).

O'Neill, S., Kreif, N., Grieve, R., Sutton, M., Sekhon, J.S., 2016. Estimating causal effects: considering three alternatives to difference-in-differences estimation. Health Serv. Outcomes Res. Methodol. 16, 1–21. https://doi.org/10.1007/s10742-016-0146-8.

Picard, R.R., Cook, R.D., 1984. Cross-validation of regression models. J. Am. Stat. Assoc. 79 (387), 575–583. https://doi.org/10.2307/2288403.

Pohl, P., Sandlund, M., Ahlgren, C., Bergvall-Kåreborn, B., Lundin-Olsson, L., Wikman, A.M., 2015. Fall risk awareness and safety precautions taken by older community-dwelling women and men—A qualitative study using focus group discussions. PLoS One 10 (3), e0119630.

Ryén, L., Svensson, M., 2015. The willingness to pay for a quality adjusted life year: a review of the empirical literature. Health Econ. 24 (10), 1289–1301.

Schepers, P., den Brinker, B., Methorst, R., Helbich, M., 2017. Pedestrian falls: a review of the literature and future research directions. J. Saf. Res. 62, 227–234. https://doi.org/10.1016/j.jsr.2017.06.020.

Sills, E.O., Herrera, D., Kirkpatrick Jr., A.J., A. B, Dickson, R., Hall, S., et al., 2015. Estimating the impacts of local policy innovation: the synthetic control method applied to tropical deforestation. PLoS One 10 (7), e0132590. https://doi.org/10.1371/journal.pone.0132590.

The Swedish Transport Administration, 2018. Analysmetod och samhällsekonomiska kalkylvärden för transportsektorn (No. 6.1). Retrieved from https://www.trafikverket.se/contentassets/4b1c1005597d47bda386d81dd3444b24/asek-6.1/asek_6_1_hela_rapporten_180412.pdf [2019-08-16]. .