SUPER SPIKES:
THE LATEST CONTROVERSY IN RUNNING SHOE TECHNOLOGY

ADAM VISOKAY§

ABSTRACT. I estimate the impact of a new advancement in running shoe technology (”super spikes”) on athletic performance in track and field competition. I exploit exogenous variation in the regulation of super spikes which approved shoes designed for distance runners but not for sprinters. I use repeated cross sectional data and a difference-in-difference approach to produce a population level estimate of how access to super spikes improves the performance of NCAA runners in 2021. I find evidence that super spikes provide a 0.45% annual improvement in performance.

1. Introduction

Eliud Kipchoge made running history as part of the breaking 2 project where he attempted to run under two hours for the marathon at Italy’s Monza race track on May 6, 2017. Though he did not break the barrier, his 2:00.25 official mark was the fastest time ever run in the history of the marathon. In addition to rewriting the history books, this feat ushered in a new era of running shoe technology. Eliud and his fleet of pace makers each wore a pair of prototypical racing shoes, characterized by their light weight carbon plates and thick, highly responsive foam. These racing shoes were a chunky departure from the sleeker models worn by elite athletes at the time. Anecdotal evidence that these ”super shoes” improved performance mounted with time, as more elite athletes and casual runners alike gained access to this new technology and set new personal

---

Date: November 20, 2022.
I thank Tyler McCormick and the rest of the study group community at UW for their attention and feedback on this project. I also thank my running friends Garrett Heath and Tripp Hurt for discussing the ideas explored in this paper on our rainy runs around Green Lake.
§ University of Washington, Center for Statistics and the Social Sciences.

bests. In light of these technological developments and subsequent allegations of unfair advantage, an investigation by World Athletics, the governing body for track and field, prompted a new suite of standardized regulations aimed at protecting the integrity of the sport. Athletes were required to use footwear compliant with these new specifications in order to record official marks in World Athletics sanctioned competition.

Are these shoes really "super", or is it just clever marketing? Several studies by kinesiologists and biomechanical engineers have explored this question in recent years. Joubert and Jones (2021) show that this carbon plated super shoe technology does improve running economy[2]. And while running economy and running performance are positively correlated for many athletes, the exact relationship remains unclear. Bermon et. al (2021) consider running performance more directly by looking at the official results from several elite road races from 2016-2019. They find that both men and women ran faster on average in years after the super shoes were adopted, but controlling for covariates that may also contribute to performance was not possible given the limitations of their data[3]. To the best of my knowledge, my paper is the first to directly measure the causal impact of new shoe technology on performance at a population level.

In order to identify a causal relationship in this context I rely upon exogenous variation in regulation of super shoe technology as it was adapted from road shoes to track and field specific footwear. The primary difference between shoes designed for the road and shoes designed for the track has to do with the competition surface. Track shoes are often called spikes because of their characteristic cleat-like needles which are designed to improve friction between the shoe and the vulcanized rubber surface of the track. Because of the apparent performance boost attributable to the super shoes used in road racing since 2017, manufacturers began adapting super shoe technology to track and field spikes. It is important to note that unlike in road racing, where runners would be considered either middle or long distance runners, track and field includes sprinters in addition to middle and long distance runners. The technical differences between sprinter spikes and distance spikes are beyond the scope of this paper, but suffice to say that manufacturers developed a separate sprinter super spike and distance super spike, both incorporating some variation of the carbon plate and responsive foam seen in the super shoes.

The first super spike prototypes were rumored to be unveiled at the 2020 Olympic Games in Tokyo, Japan by a select group of elite track and field athletes before mass manufacturing and availability in 2021. However, in anticipation of how super spikes could affect track and field performance at the Olympic Games, World Athletics announced a new regulatory framework to govern the adoption super spikes. Interestingly, this regulation by World Athletics ended up certifying the distance super spikes but not certifying the sprinter super spikes. The controversy surrounding this development was drawn out for another year when the Tokyo organizers announced the Games were to be postponed due to the COVID-19 pandemic.

Because of the timing of the regulation and the pandemic, super spikes never had their breaking 2 moment like their road shoe predecessors. Instead, manufacturers decided to scrap their production of the sanctioned sprinting super spikes but continued their production of distance super spikes as planned in 2021. The timing of this new regulation served as an exogenous shock to super spike access whereby distance super spikes were available and sprinter super spikes were not.

The remainder of this paper is as follows: in Section 2 I will describe the data I collected and share some summary statistics. Section 3 is devoted to my identification methodology. In Section 4 I share my primary results. I conclude in Section 5 with a discussion of the key limitations of this exercise and suggested future work.

2. Data

The population of interest in this paper is that of student-athletic-workers competing in Outdoor NCAA track and field in the United States. I choose the designation student-athletic-worker (SAW) to reflect the labor these athletes provide to the academic institutions which they represent. This population of NCAA SAW's is well defined and competition results are consistently recorded and reported by the Track and Field Results Reporting System (TFRRS), the source of the data for this project. I scraped the top 100 regular season individual SAW performances for each running event, for both men and women, for all three NCAA Divisions DI, DII, DIII, and for the years 2010 to 2021 with the exception of 2020 due to the competition cancellations caused by the COVID-19 pandemic.

In total these data include 79,200 unique observations (n=79200). For each observation I observe the athlete's name, their school affiliation and associated NCAA Division, their sex, the date and location of competition, their event and their official result. The sprinting events include the 100m, 4x100m, 100m hurdles (for women), 110m hurdles (for men), 200m, 400m, 4x400m and 400m hurdles. The distance events include the 800m, 1500m, 3000m steeplechase, 5000m and 10000m.

It is important to note that I do not directly observe the choice of footwear for each SAW's performance in these data. I rely upon the assumption that because of the wide spread access and apparent competitive advantage of distance super spikes, a vast majority of the top marks in distance events in 2021 were achieved by SAW's wearing super spikes. In other words, I implicitly assume all distance performances in 2021 were aided by super spikes, but not in any other years or for sprinting events in 2021. More on this important assumption in Section 5.

Throughout the remainder of this paper I will refer to each subgroup permutation of NCAA division-sex-event as a cohort. That being said, some but not all SAW's may or may not be represented in multiple divisions, sexes, events in different years. Thus cohort in this sense refers specifically to the aggregate rather than individual group characteristics. For example, Figure 1 describes the distribution of performances for the Division I-Men's-1500m cohorts. This demonstrates both that the distribution within a year appears to be skewed towards faster performances and that the distribution across years appears to be getting faster over time.

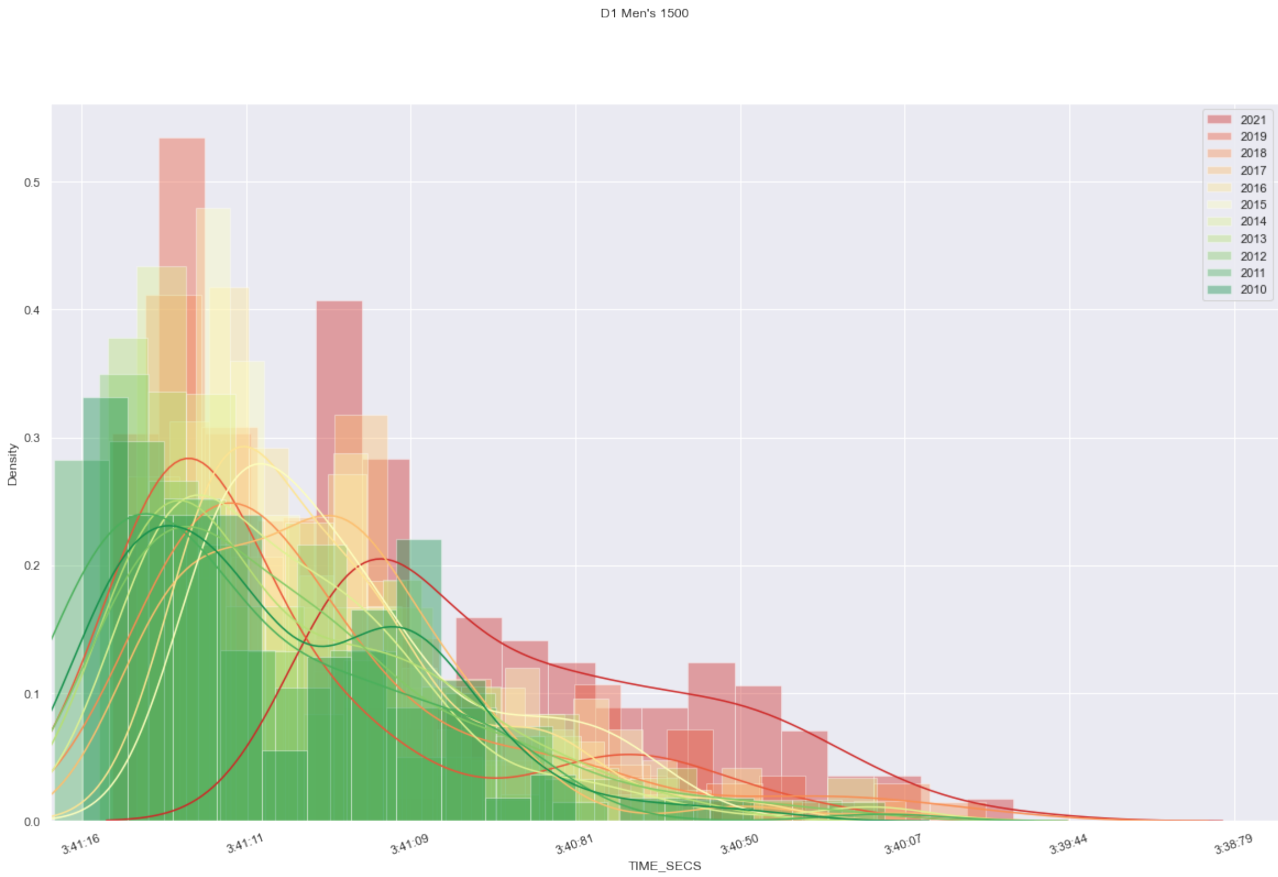

FIGURE 1. Division I Men 1500m Performance Distributions

To look more closely at the change in average performance for a different cohort, the Division I-Men's-100m for example, I compute the average performance within each year and plot the results in Figure 2. A line of best fit for the years 2010-2019 is shown in blue and the mean performance in 2021 is shown in red. Note negative the slope, indicating that average performance is trending faster with time for this cohort.

Now, to consider the trend in performance from 2010-2019 for all cohorts simultaneously, I first estimate the percentage change in performance on average for each cohort in the sample and then plot the distribution of those dis-aggregated into sprinting and distance events in Figure ??. This exercise suggests that SAW's are performing a fraction of a percentage point faster each year on

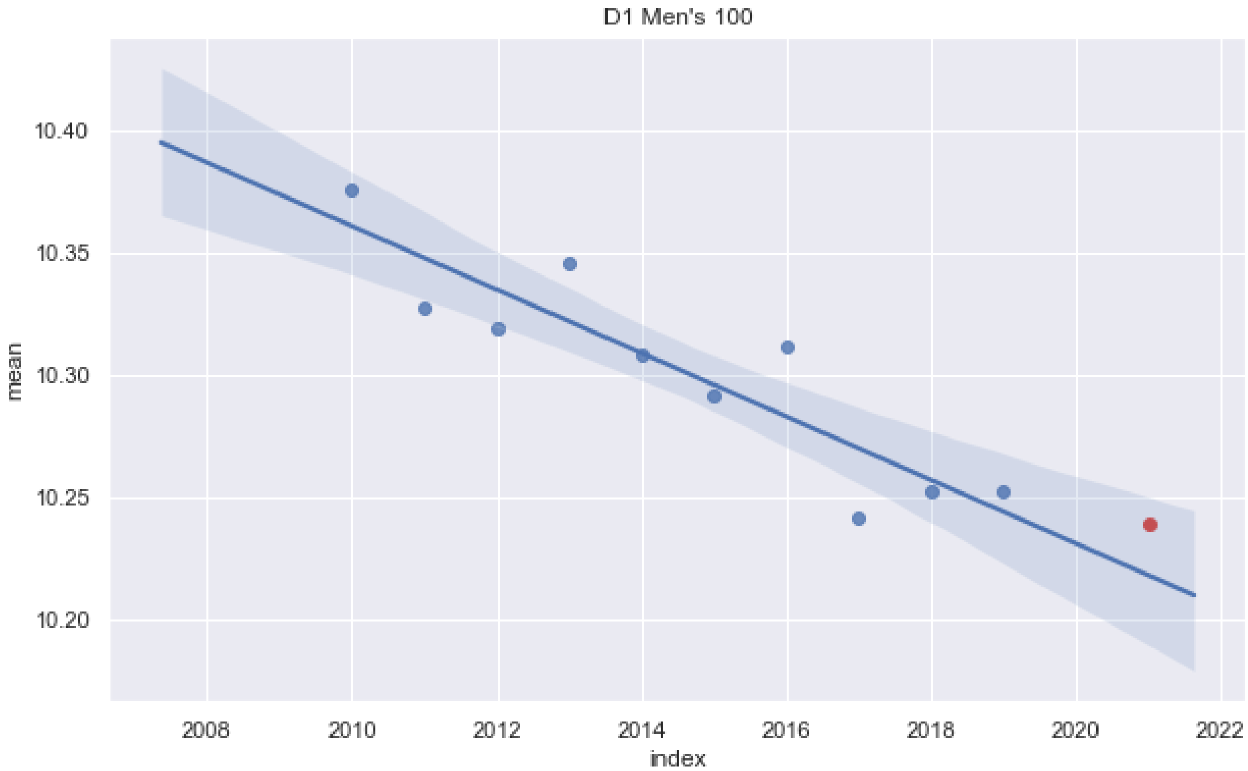

FIGURE 2. Division I Men 1500m Trend in Improvement

average over this period (sprinters and distance runners improving by 0.13% and 0.11% on average, respectively.)

3. Methodology

I wish to evaluate the causal effect of super spikes on running performance over the population NCAA Track and Field SAW's using data from 2010-2021. In order to compare running performance results across cohorts I created two synthetic groups, one each for distance events and sprinting events. I took the average performance each year from 2010-2019, then indexed using 2010 as the base year, and then took an unweighted average across cohorts to create separate synthetic distance and sprinter groups.

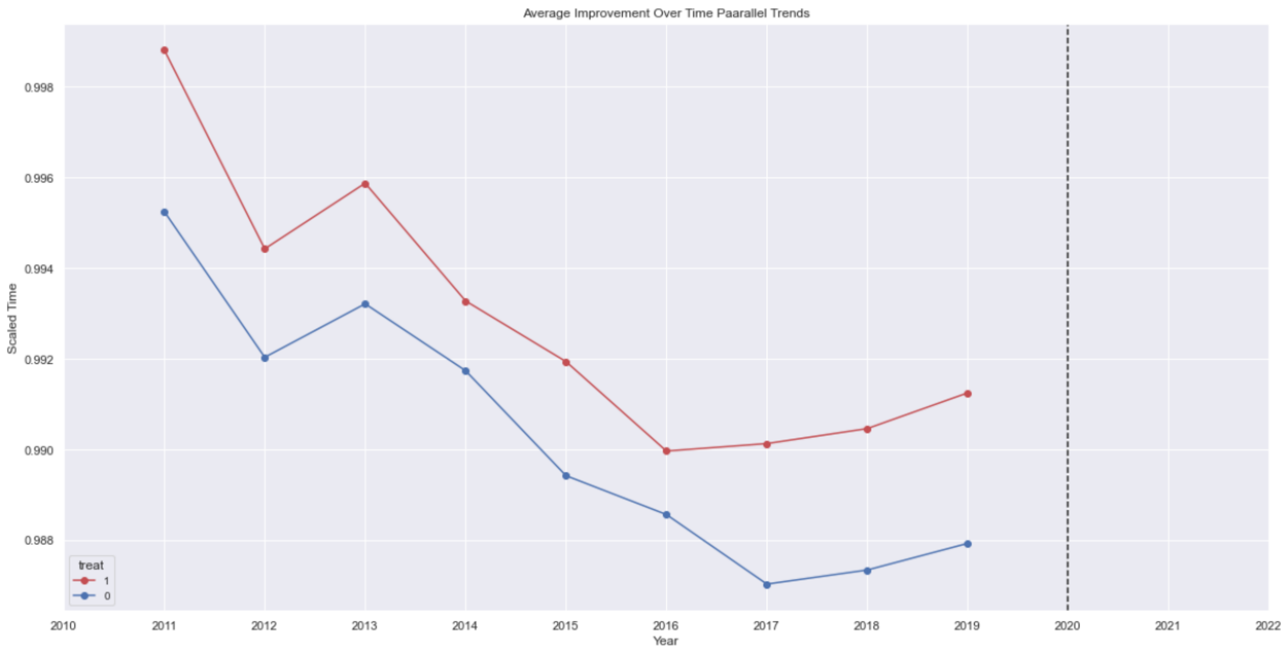

Figure 3. Parallel Trends for Synthetic Distance and Sprinter Groups

A primary challenge in causal inference is controlling for all of the possible covariates that may impact the outcome. In the context of running, there are many variables which contribute to performance. Some covariates, such as nutrition, body composition, and shoe technology, can be directly observed. But there are also latent variables which contribute to performance, including self confidence and risk aversion, which cannot be directly observed. Most of the relevant determinants of running performance are not available in the TFRRS data described in Section 2. In order to overcome this challenge and estimate a causal effect of super spikes on running performance, I use a DiD approach described in Card and Krueger (1993)[1]. This approach relies upon the assumption that the control and treated groups exhibit parallel trends in outcome before treatment, in essence controlling for all shared covariates between the two groups and allowing for the isolation of the treatment effect. Figure 3 demonstrates the parallel trend in running performance between sprinter and distance groups ahead of 2021. The specification for this particular DiD is as follows:

$$\Delta P_i = \alpha + \beta Treat_i + \gamma After_i + \delta Treat_i * After_i + \epsilon \tag{1}$$

where ΔP_i is the change in running performance, α is a constant term, β is running event effect, γ is a championship year effect, and δ is the true treatment effect for distance events in 2021.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:       TIME_SECS_SCALED   R-squared:                       0.035
Model:                            OLS   Adj. R-squared:                  0.031
Method:                 Least Squares   F-statistic:                     9.388
Date:                Tue, 12 Oct 2021   Prob (F-statistic):           4.22e-06
Time:                        22:13:35   Log-Likelihood:                 2849.2
No. Observations:                 792   AIC:                            -5690.
Df Residuals:                     788   BIC:                            -5672.
Df Model:                           3
Covariance Type:            nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
Intercept         0.9913      0.000   3057.093      0.000       0.991       0.992
AFTER[T.True]     0.0002      0.001      0.187      0.852      -0.002       0.002
TREAT             0.0024      0.001      4.702      0.000       0.001       0.003
TREAT_AFTER      -0.0045      0.002     -2.715      0.007      -0.008      -0.001
==============================================================================
Omnibus:                       22.417   Durbin-Watson:                   0.725
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               24.431
Skew:                          -0.373   Prob(JB):                     4.95e-06
Kurtosis:                       3.429   Cond. No.                         8.66
==============================================================================
```

FIGURE 4. Regression Results

4. CONCLUSION

The ordinary least squares regression results specified in Equation 1 are shown in Figure 4. Figure 5 shows a visualization of this average treatment effect as the vertical distance between the

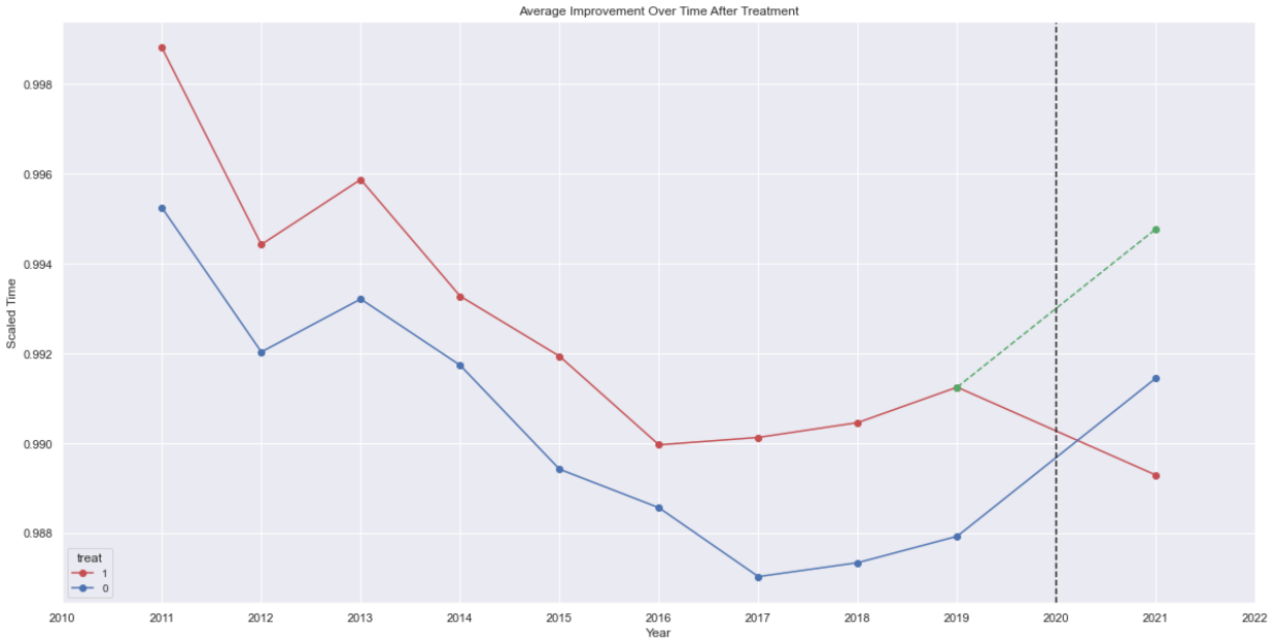

Figure 5. Treatment Effect

true treated group in red and the counterfactual distance runner group in the absence of treatment in green.

The synthetic sprinter group performed worse in 2021 than any of the previous 5 years of available data, while the distance runner group performed at an all time best. Assuming there are no systematically different effects impacting performance of distance runners as compared to sprinters other than access to super spikes, the 0.45% improvement compared to the counterfactual result observed in distance runner performance in 2021 is attributable to super spikes.

Recall the discussion from Section 2 where the annual improvement in performance was between 0.11%-0.13% for distance runners and sprinters, respectively. This effect size is more than three times larger! For real world context, consider some of the resent real world results from the 2022 Toyko Olympic Games. Sifan Hassan delivered victory for the Netherlands the Womens 5000m in a time of 14:36.79, followed by Helen Obiri in 14:38.36, Gudaf Tsegay in 14:38.87 and Agnes Tirop in 14:39.62. 0.45% of Hassan's winning mark is 3.95 seconds, more than enough to be the difference between just missing the podium and winning gold in Toyko that Summer.

References

[1] David Card and Alan B Krueger. *Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania.* Working Paper 4509. National Bureau of Economic Research, Oct. 1993. DOI: 10.3386/w4509. URL: http://www.nber.org/papers/w4509.

[2] Dustin P. Joubert and Garrett P Jones. "A Comparison of Running Economy Across Seven Carbon-Plated Racing Shoes." In: *Kinesiology and Health Science* Faculty Publications.33 (2021). URL: https://scholarworks.sfasu.edu/kinesiology/33?utm_source=scholarworks.sfasu.edu%2Fkinesiology%2F33&utm_medium=PDF&utm_campaign=PDFCoverPages.

[3] Bermon S. et al. "Effect of Advanced Shoe Technology on the Evolution of Road Race Times in Male and Female Elite Runners." In: *Front. Sports Act. Living.* 3:653173 (2021). DOI: 10.3389/fspor.2021.653173.