Using weather forecasts to forecast whether bikes are used

by
Jan Wessel

Institute of Transport Economics Münster
Working Paper
No. 32
June 2020

wissen.leben
WWU Münster

Address
Institut für Verkehrswissenschaft
Am Stadtgraben 9
D 48143 Münster, Germany

Telephone
+49 251 83-22 99 4

Fax
+49 251 83-28 39 5

E-Mail
verkehrswissenschaft@uni-muenster.de

Website
http://www.iv-muenster.de

The Working Paper Series seeks to disseminate economic research work by the WWU, Institute of Transport Economics staff and visitors. Papers by researchers not affiliated with the WWU Institute of Transport Economics may also be considered for publication to the extent that they have been presented at research seminars/workshops organised by the institute.

The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors' own opinions.

The Series is managed by the Director of the Institute of Transport Economics.

Using weather forecast to forecast whether bikes are used

By Jan Wessel*

Although several papers have shown that bike ridership is affected by actual weather conditions, this is the first study to comprehensively investigate the impact of forecasted weather conditions on bike ridership. The results show that both actual and forecasted weather conditions can be used as useful explanatory variables for predicting bicycle usage. Even incorrect weather forecasts can impact on bike ridership, which underlines the importance of weather forecast effects for traffic planners; for example, forecasted rain can reduce bike traffic by 3.6 % in periods that turn out to be rain-free. Additionally, a digital image-processing method is used to calculate the darkness of the cloud coverage displayed on weather forecast maps. The results imply that bike ridership is significantly smaller in regions with darker forecasted clouds. It is also shown that weather forecasts have a stronger impact on recreational bike traffic than on utilitarian traffic. Furthermore, various lagging and leading effects of rain forecasts are outlined. Morning rain forecasts can, for example, reduce bike ridership in midday and afternoon hours that were predicted to be rain-free. To derive these results, hourly bicycle counts from 188 automated counting stations in Germany are collected for the years 2017 and 2018. They are linked to actual weather data from Germany's National Meteorological Service and with historical weather forecasts that are deduced from weather maps of Germany's most-watched television news program. Log-linear and negative binomial regression models are then used to estimate the weather forecast effects.

JEL: R49.

Keywords: Cycling, bike ridership, automated counting stations, weather conditions, weather forecasts, image processing.

I. Introduction

Traveling by bike can have many positive effects, both for oneself as well as for others. It can not only provide substantial health benefits for the cyclist (de Hartog et al., 2010), but also contribute to reducing pollution and congestion (Krizek, 2007). Thus, the promotion of cycling is strongly recommended in order to promote a healthier and less-polluted environment (Handy et al., 2002). One important aspect that needs to be taken into consideration when designing policies and infrastructure for increasing the share of bike trips, is the effect of weather on biking. Several papers (e.g. Miranda-Moreno and Nosal, 2011) have shown that poor weather conditions such as rain can significantly reduce bike ridership.

While several papers found that actual weather conditions have a significant impact on bike ridership (e.g. Miranda-Moreno and Nosal, 2011; Nosal and Miranda-

* Westfälische Wilhelms-Universität Münster, Institute of Transport Economics, Am Stadtgraben 9, 48143 Münster, Germany.

The final publication in "Transportation Research Part A: Policy and Practice" is available online at: https://doi.org/10.1016/j.tra.2020.06.006

Moreno, 2014), the effects of weather forecasts – and, more importantly, of wrong weather forecasts – on bike ridership have so far been neglected in the literature. Since using the bike has ramifications for many aspects of the daily schedule (e.g. earlier wake-up times due to longer commuting times), the decision to use the bike is often made, or even has to be made, the day before the actual trip. Consequently, it is reasonable to assume that weather forecasts significantly affect bike ridership.

Meng et al. (2016) conduct a survey to confirm this notion. They find that 66.5 % of their participants would change their travel behavior if rain was forecasted. Moreover, Kraemer et al. (2015) use rush-hour bicycle counts to show that bike ridership increases with higher forecasted air temperatures and that it decreases with a higher forecasted chance of rain. While these two papers highlight the importance of weather forecasts for bicyclists, a more comprehensive analysis of different weather forecast effects on bicycle ridership is provided within this research article. Using data from automated bicycle counting stations and a larger set of weather forecast variables, the weather impacts are estimated (i) using only actual weather data, (ii) only forecasted weather data, and (iii) using a combination of both actual and forecasted weather data. In the combined regression model, we estimate effects of wrong weather forecasts, which have a significant impact on bike ridership. Moreover, a novel digital image-processing method is introduced to evaluate the effect of the cloud coverage displayed on weather forecast maps. It is also shown that weather and weather forecast effects depend crucially on the type of bike traffic (utilitarian, recreational, or mixed). As recreational bike trips are often easier to postpone or cancel, they are more affected by actual weather and weather forecast than utilitarian bike trips. Additionally, lagging and leading effects of weather forecasts are estimated. The analysis of these effects is important as, for example, rain forecasts for morning hours can deter people from going to work by bike, subsequently also reducing bike ridership in afternoon hours when these people would commute back home.

The novel results of this study underline that not only actual weather conditions, but also weather forecasts should to be considered in order to correctly predict bike traffic. Moreover, wrong weather forecasts can lead to suboptimal levels of bike ridership. This is especially important for cases of wrongly forecasted rain, because people are likely to switch to less healthy and less environmentally-friendly modes of transport. The results are thus helpful for traffic planners and policymakers who should consider these effects in planning and implementing strategies to increase bike ridership.

The remainder of this paper is organized as follows. Section II reviews the literature on the relationship between biking and weather conditions in general, as well as the scarce literature on weather-forecast effects. The data and descriptive statistics are outlined in Section III, and Section IV introduces the methodology. The results are presented in Section V. Section VI discusses the results and concludes.

II. Literature review

Within this section, we first review papers dealing with the impact of actual weather conditions on bicycle ridership. These papers are divided into two groups, based on the type of data that is used. For a more comprehensive review of papers dealing with the impact of actual weather conditions on bicycle traffic and also on other modes, we refer to Böcker et al. (2013). In a second step, we look at the few papers that touch at the subject of weather forecasts and their impact on travel behavior. In a third step, gaps in research are derived and it is outlined how they are addressed in this paper.

A. The impact of actual weather conditions (survey data and travel log data)

Hanson and Hanson (1977) use travel diary data from Uppsala (Sweden) and find that air temperature has a positive impact on the share of trips made by bicycle, whereas cloud coverage has a negative impact. They also show that commuters are less sensitive to changes in air temperature than non-commuters are. In accordance with these findings, Helbich et al. (2014) show for the Greater Rotterdam area in the Netherlands that cycling trips are positively affected by air temperature, and negatively by wind speed and precipitation. They also find that the weather effects vary across different regions, and that weather conditions have a smaller impact on bicycle trips in densely settled areas than in low-density areas.

For Sweden, Liu et al. (2015) use survey data to show that the effect of weather on bicycle ridership and also the subjective perceptions of weather conditions are not constant, but can differ between regions and between seasons. They show, for example, that people who live in generally colder regions are more aware of seasonally unusual air temperature variations during spring and autumn seasons than people who live in generally warmer regions. During winter seasons, however, the people from colder regions are less aware of such air temperature variations due to very low average air temperatures. For a similar setting, Liu et al. (2014) analyze weather effects on non-commuter trips, and differentiate between routine activities (e.g. daily shopping) and leisure activities (e.g. visiting friends). They find that routine and leisure trips can be influenced differently by weather conditions and that there are again regional differences in these effects. They argue that it is important to account for regional idiosyncrasies, as weather conditions can even have two opposing effects in two different regions.

Also, many papers find that social characteristics such as gender, age, biking experience, or occupational status, have a significant impact on individual perceptions of weather conditions and their reactions to adverse conditions (Nankervis, 1999; Bergström and Magnusson, 2003; Winters et al., 2007; Heinen et al., 2011; Saneinejad et al., 2012; Flynn et al., 2012; Ahmed et al., 2013; Motoaki and Daziano, 2015, e.g.). Zhao et al. (2018b) show for Beijing that bicyclists not only respond to weather conditions, but also to pollution. The generally negative effect of $PM_{2.5}$ concentration appears to be higher for females and for individuals with medium or high income, who would subsequently switch to other modes of transport.

B. The impact of actual weather conditions (bicycle counting data and bike-share data)

In addition to papers using survey or travel log data, there are various papers that analyze the impact of actual weather conditions on bicycle counts. Such bicycle counts can be measured either by hand or automatically, and they provide a good indicator of actual bike ridership at a specific location. Thomas et al. (2013) use daily bicycle counts from 16 cycle paths in two cities in the Netherlands. They show that average daily air temperature and sunshine duration have a positive impact on daily bicycle counts, whereas higher precipitation and wind velocity reduce daily bike ridership. Weather has a significantly larger impact on recreational bike traffic than on utilitarian bike traffic. The authors also show that 80 % of daily bicycle count variations are caused by weather conditions. Ermagun et al. (2018) also analyze the effect of weather conditions on daily bicycle counts. They include data from thirteen cities in the U.S. and find differences relating to responses to weather, both within and across regions.

Miranda-Moreno and Nosal (2011) analyze bicycle counts per hour for five automated counting stations in Montreal (Canada). Using both an absolute and a relative ridership regression model, they estimate the effects of air temperature, precipitation, wind-speed, and relative humidity on hourly cycle volumes. They find that absolute ridership increases by 4-5 % if the air temperature increases by 10 %, relative to the overall average air temperature. An increase in air temperature, however, can decrease bicycle counts if the air temperature is above 28C and relative humidity above 60 %. Moderate or severe rain, in combination with fog, drizzle, or freezing rain, reduces ridership by 19 %. There is also a negative effect of lagging variables. Rain in the past three hours causes ridership to drop by 25 to 36 %, and rain in the morning reduces ridership in rain-free afternoon hours by 13 to 15 % if all other factors are held constant. The relative ridership model finds that a 1 % increase in average air temperature increases hourly ridership by 2.5 %. The negative effect of humidity appears weaker in this analysis, but is still significant. Nosal and Miranda-Moreno (2014) conduct a similar analysis and include both recreational and utilitarian bicycle counting stations. They find non-linear effects of air temperature and humidity, and also observe that recreational bike traffic is more sensitive to weather conditions, as did Thomas et al. (2013) for daily bicycle counts.

Zhao et al. (2018a) use data from two automated bicycle counters in Seattle (USA) to show that the traffic measured on an on-road bike lane, which represents a utilitarian cycling pattern, reacts differently to weather than traffic on an off-road bike trail, which represents a recreational cycling pattern. The weather is less likely to affect weekend cycling on the trail than weekend cycling on the on-road bike lane, whereas the opposite is true for weekday cycling. Similar to Miranda-Moreno and Nosal (2011), they also find that rain has a lagged negative impact. Miranda-Moreno et al. (2013) state that local climate appears to influence monthly ridership profiles across cities in North America, and Tin Tin et al. (2012) also find that weather influences hour-to-hour and day-to-day variations in bicycle ridership.

Besides data from bicycle counting stations, weather effects can also be estimated with data from public bike-sharing programs. An et al. (2019) analyze data from such a program in New York (USA) and show that weather has a stronger impact on cycling rates than infrastructure, topography, or calendar events. For Brisbane (Australia), Corcoran et al. (2014) show that windspeed and rainfall exert a significant negative effect on the number of bike-sharing trips that are made. Air temperature, on the other hand, has a positive, yet smaller effect on bike-sharing trips. Moreover, Gebhart and Noland (2014) use bike-share data from Washington, DC (USA) to show that not only the number of trips, but also their duration decreases under adverse weather conditions such as low air temperatures or rain. De Chardon et al. (2017) show, for regions with temperate or continental climate, that the number of daily bicycle-sharing trips decreases in winter months, whereas these reductions are less pronounced in regions with warmer climates. In addition to these findings, the impact of climate change on bikeshare usage is analyzed by Heaney et al. (2019). In a first step, they estimate the relationship between air temperature and bikeshare usage. They then predict future bikeshare usage, depending on long-term air temperature predictions from different climate models. They conclude that the higher air temperatures associated with climate change could increase bikeshare usage by up to 3.1 % by 2070.

C. *The impact of weather forecasts*

While there is a considerable literature on the effects of actual weather conditions on bike ridership, the interplay between weather forecasts and bike ridership has thus far been barely studied. Böcker et al. (2013, p. 86) point out that "a detailed exploration of the role of expectations and weather forecasts for travel or activity behavior is yet to be achieved".

There are, however, a few papers that look explicitly at weather forecasts and biking. The first is by Meng et al. (2016), who conduct a survey with 553 cyclists in Singapore. Of the questionnaires, 223 were collected during wet weather and 330 during good weather. Among other questions, the survey participants were specifically asked whether they had obtained weather forecast information before the trip, and if so, where they obtained this information (e.g. internet, television, smart phone apps, etc.). Moreover, the participants were asked whether they would change their travel plan if the weather forecast predicted bad weather. The results show that in Singapore, 28.2 % of survey participants acquired information from weather forecasts before their trip. People under 30 years, male bicyclists, and employed bicyclists appear most likely to acquire weather forecast information before their trip. Moreover, if individuals are less certain how the weather will unfold, they are more likely to acquire weather forecast information. Another interesting result is that 66.5 % of all survey participants stated that they would change their travel plan if rain were forecasted. For cyclists who checked weather forecasts before the trip, this percentage is higher (69.3 %), while it is lower for cyclists who did not check weather forecasts (34.1 %). Also, the analysis shows that cyclists who traveled to work or school, would likely transfer to other modes of transport if rain were forecasted, whereas leisure cyclists would postpone their trip.

The second paper is by Kraemer et al. (2015), who counted the number of bicycle commuters in Washington, DC (USA), in 96 hourly sessions during rush-hour periods. As predictive variables, they included both actual and forecasted weather conditions (predicted high air temperature, rain chance, actual rain). They show that the predicted air temperature can increase bike ridership by 2.2 % per F. Furthermore, if the weather forecast predicts that the chance of rain is above 50 %, the number of bicyclists will decrease by 40 %, and actual rain will decreases the number of bicyclists by 28 %. They do not, however, elaborate on whether there are interaction effects between the actual weather and weather forecast.

Additional insights into the influence of weather forecasts on activity-travel behavior are provided by Cools and Creemers (2013), who conduct a stated adaption experiment and show that the forecasted weather conditions have a significant effect on the likelihood of changing activity-travel behavior. Especially snow and temperatures above 28C can trigger such changes. It is interesting to note that the exposure to weather forecasts, the specific media source, or the perceived reliability of the forecast have no significant effect on the probability of changing activity-travel behavior.

Another paper that touches on the relationship between weather forecasts and biking is from Dias et al. (2015). Although their paper focuses on the technical prediction of the occupancy status of bicycle storing facilities in Barcelona, they state that three-day weather forecasts can improve the quality of their prediction models. This underlines the importance of further studying the relationship between weather forecasts and bike ridership.

D. *Gaps in research*

While the aforementioned papers have analyzed selected relationships between weather forecasts and cycling behavior, there are still some important gaps in research.

First, the previous analyses have used only very few and rather crude forecast variables. Kraemer et al. (2015), for example, use the daily predicted high air temperature and the chance of rain as forecast variables. In our study, however, we additionally estimate the impact of various other weather forecast variables on bicycle ridership. Our set of weather forecast variables includes the forecasted air temperature as well as dummy variables on whether light clouds, dark clouds, rain, thunderstorms, snow, or weather warnings are forecasted. Moreover, a novel digital image-processing method is introduced to evaluate the impact of the cloud coverage displayed on weather forecast maps. This extensive set of weather forecast variables generates additional insights on how different weather forecast variables can influence bike ridership.

Second, it has not been studied how wrong weather forecasts can influence bike ridership. It can be assumed that the decision to use a bike is not always made right before the trip, but already in advance. Thus, people might decide to not use their bike for their next-day morning commute if there is a bad weather forecast for that day. Consequently, they make arrangements to get to work with a different means of transport (e.g. reserving a shared car or organizing a carpool). If such arrangements are fixed, it would be more difficult to spontaneously change their plans on the next morning in cases of actually good weather. Thus, wrong rain forecasts could lead to a reduction in the number of bike trips on the next day. In this paper, we estimate the impact of wrong rain forecasts, as well as the impact of wrong good-weather forecasts on bicycle ridership to better understand how wrong weather forecasts can influence actual bike trips.

Third, it is not studied in the literature if the effects of weather forecasts differ between recreational and utilitarian bicycle traffic. It could be argued that recreational bicycle trips are easier to postpone or to cancel than utilitarian bike trips, and thus more sensitive to adverse weather forecasts. By comparing the effects of different weather forecast variables on bike ridership at recreational, utilitarian, and mixed bicycle counting stations, we can provide valuable information on how different traffic types are affected by weather forecasts.

Fourth, the literature has not looked at lagging and leading effects of weather forecasts. Commuters might, for example, resort to not use their bike if rain is forecasted for either the morning or the afternoon commute. Therefore, a rain forecast for morning hours could have a lagging impact on bike ridership in afternoon hours, and a rain forecast for afternoon hours could have a leading impact on bike ridership in morning hours. Such time-shifted effects are estimated with the help of various dummy variables and provide additional insights on how the behavior of bicyclists can be influenced by weather forecasts.

III. Data and descriptive statistics

A. Data

Bicycle counter data

We acquired hourly bicycle count data from 188 bicycle counting stations from 37 different cities and regions all over Germany.[1] 140 of these 188 stations offer hourly bicycle counts for the whole sample period, that is from 01.01.2017 to 31.12.2018, and 175 stations offer data for at least 365 consecutive days. On average, a bicycle counting station of the sample offers hourly data for 668.2 days (91.5 % of the total sample period). For our regressions, we use data from all 188 stations, as the missing observations appear to be randomly distributed over the total sample period.

All bicycle counting stations in the sample were installed by Eco-Counter, a company specialized in pedestrian and bicycle counting solutions. The raw data, that is the number of bicycles per hour, are the property of the respective cities and regions, but were provided to us by the (bicycle) traffic planners or other contact persons. As the cities and regions decide where to install the counting stations, they are distributed rather randomly across different locations (e.g. separated bicycle lanes, multi-use paths, bicycle tracks,...). The accuracy of the counting stations is around 95 %.

An overview of the regions and cities in the sample, their locations, and the number of stations per city or region can be found in Figure 1.

Actual weather data

Various factors can impact on bicycle counts. Among the most important of these is the actual weather. Hourly weather data is provided by Germany's National Meteorological Service (Deutscher Wetterdienst) and includes the air temperature (measured at 2 m above ground, in C), precipitation amount (in mm), relative humidity (in %), wind speed (measured at 10 m above the ground, in m/s), and cloud coverage (in eights).

These actual weather variables are linked to the bicycle counter data. Therefore, we always choose the weather station closest to the bicycle counting station and then combine the data. The average distance between a bicycle counting station and a weather station is 11.8 km. If we look only at the cities within our sample and disregard the regions where bicycle counting stations are more scattered, the average distance decreases to 9.4 km. As the weather stations are not directly next to the bicycle counting stations, there might be small differences between the weather conditions at weather stations and the weather conditions at counting stations. For example, a rural weather station could observe lower

[1] The cities and regions include: Augsburg, Berlin, Bochum, Bonn, Bremen, Dsseldorf, Erlangen, Essen, Freiburg, Gttingen, Hamburg, Hannover, Heidelberg, Heilbronn, Herzogenaurach, Jena, Kiel, Kirchheim unter Teck, Cologne, Leipzig, Lrrach, Ludwigsburg, Mannheim, Munich, Mnster, Nationalpark Mritz, Nuremberg, Oberhausen, Oldenburg, Rhein-Erft-Kreis, Rhein-Kreis-Neuss, Rhein-Sieg-Kreis, Rostock, Stuttgart, Sauerland Radwelt, Tbingen, Wrzburg. All data belong to the respective regions. It is greatly appreciated that the company Eco-Counter helped in acquiring the data.

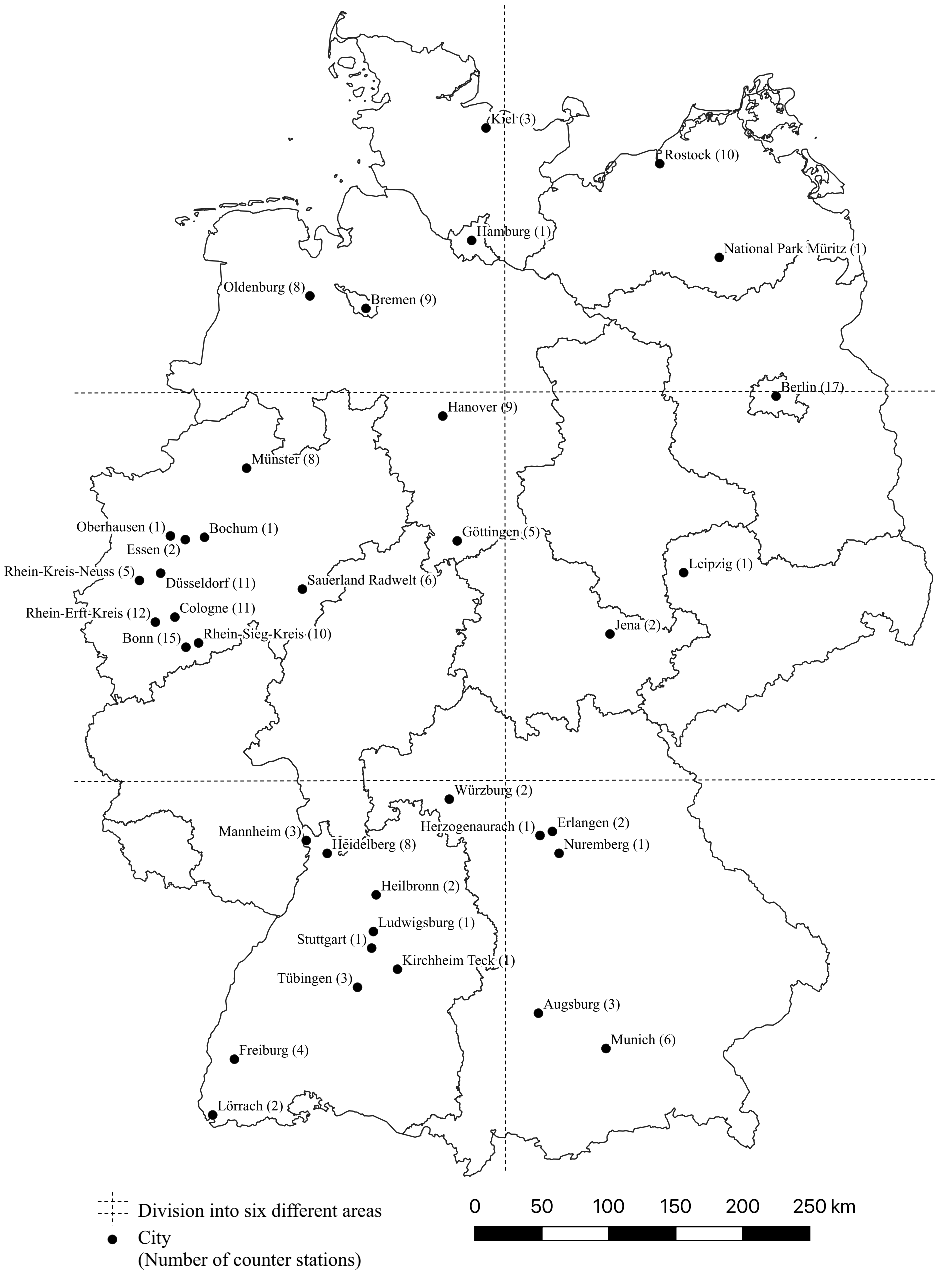


Figure 1. : Cities and stations of the sample

temperatures than the city-located bicycle counting station. If such differences are systematic, however, they can be largely accounted for by using station fixed effects. Moreover, the high quality and comparability of the weather station data from Germany's National Meteorological Service is a considerable advantage over including data from different weather data providers that might operate weather stations which are slightly closer to the counting stations.

Weather forecast data

Hourly bicycle ridership is not only determined by the actual weather conditions, but also by weather forecasts. The decision to commute to work either by bike or using a different mode of transport, is often not decided in the morning right before leaving the house, but might have already been determined the previous evening when watching the weather forecasts. Therefore, we also include different weather forecast variables in our model.

As no historical weather forecast data were available, we evaluated weather forecasts from Germany's oldest and most-watched television news program, "Die Tagesschau" (*Engl.: (Re)view of the Day*). This program is broadcast daily at 8 p.m. on German TV Channel "Das Erste", a publicly funded TV channel that can be freely received by all German residents. On a daily average, 9.63 million people watch *Die Tagesschau* (34.5 % market share).[2] The weather forecasts of *Die Tagesschau* are also used in other programs of *Das Erste* and its regional broadcasting corporations, as well as in some nightly programs of publicly broadcasted radio stations. Moreover, the forecasts are used for the TV Channel Tagesschau24, which is a publicly broadcast TV news channel, and as content for online articles and in corresponding mobile applications. Thus, it can reasonably be argued that the weather forecasts we evaluate are observed by a large share of the German population. Moreover, Cools and Creemers (2013) conducted an experiment to show that the media source of the consumed weather forecast does not significantly affect the probability of changing activity-travel behavior.

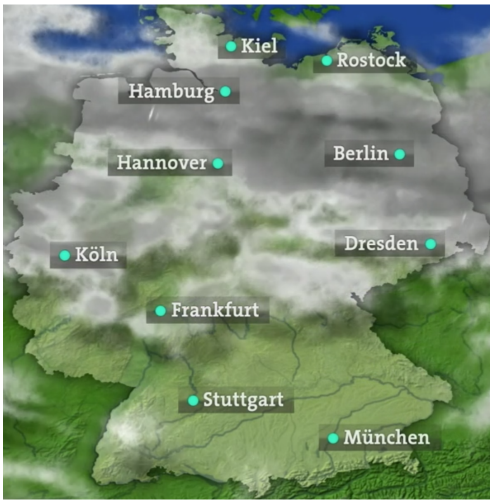
For 2017 and 2018, all broadcasts of the 8 p.m. *Tagesschau* can be accessed via the official multimedia library *ARD Mediathek* of *Das Erste*.[3] The *Tagesschau* usually runs for 15 minutes and always ends with a roughly one-minute-long weather forecast. This weather forecast consists of various animated weather maps. On the first weather map, areas of high pressure and low pressure, as well as cloud movements, are depicted for all of Europe. After this, an animated weather map shows different weather conditions (cloud coverage and movement, rain, snow, thunderstorms) for the next 24 hours in Germany, which is depicted in Figures 2a, 2b, and 2c. Also, different types of weather warnings (wind, danger of slipping, iciness, flood, and a generic warning symbol) are displayed in this weather map. Next, nightly and daily air temperatures are shown (Figure 2d). After this, weather forecasts for the three days after the main forecast are briefly outlined.

In order to manually evaluate these weather maps, we divide Germany into six different areas as displayed in Figure 1: northeast, mideast, southeast, southwest, midwest, northwest. To generate static weather maps, we use the animated weather map with the different weather conditions and extract the frames that display the weather conditions at 8:00, 12:00 and 16:00 o'clock. Next, we use the static 8:00 o'clock weather map to assign the depicted weather conditions to morning hour observations between 6:00 and 10:00 for each of the six aforementioned regions. The possible weather condition classifications are as follows: clear sky, light cloud cover, heavy cloud cover, rain, snow, thunderstorm. Furthermore,

[2] Source: `https://www.dwdl.de/zahlenzentrale/70389/tagesschau_verliert_zuschauer_aber_bleibt_deutlich_vorne/`.

[3] Only for August 19$^{\text{th}}$, 2018, is there no *Tagesschau* available. Therefore, we have to drop the next day – the day for which the weather is forecasted – from our analysis.

(a) Weather condition map 8:00



(b) Weather condition map 12:00

(c) Weather condition map 16:00

(d) Daily high air temperature map

Source: ARD / Hessischer Rundfunk.

Figure 2. : Weather conditions and daily high air temperature maps for July 4th, 2017

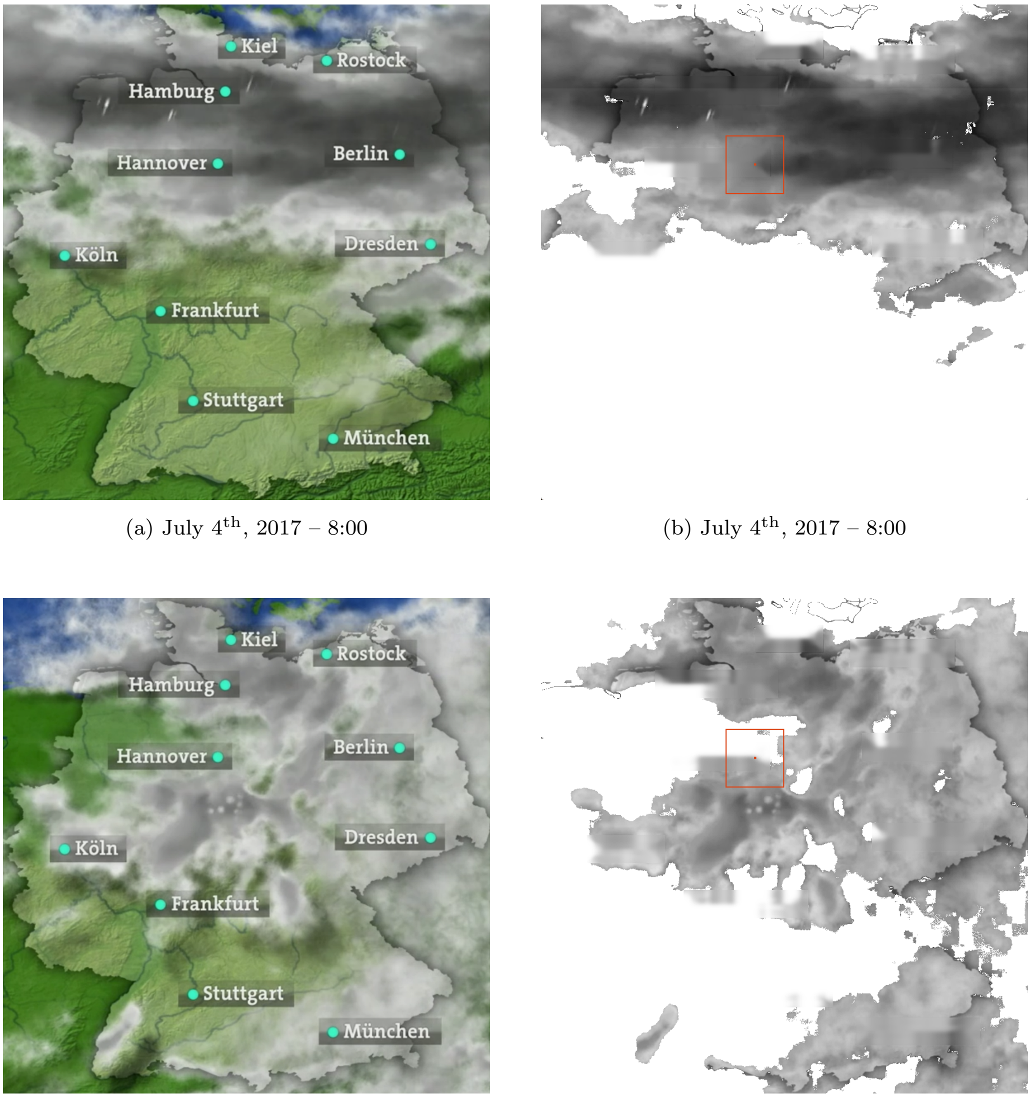
we assign the outlined weather warnings (wind, danger of slipping, iciness, flood, and a generic warning symbol) to each of the six regions. Thus we obtain weather forecast values for the morning hours of our dataset. For the midday hours from 11:00 to 13:00, as well as for the afternoon hours from 14:00 to 18:00, a similar assignment procedure is conducted with the frames of the 12:00 and the 16:00 o'clock weather maps.

It should be noted that our six weather condition classifications are not as detailed as the weather conditions provided by the Deutscher Wetterdienst (DWD), and that the actual classification of the depicted weather conditions is, to a certain degree, subjective. Moreover, as we assign only the most prominently depicted weather condition to each of the six areas, we do not account for differences between cities within one region. These drawbacks make the procedure rather crude, which is, however, exactly what we are looking for. That is, the outlined procedure thus takes into account the imperfect information intake when watching weather forecasts. This information can be misleading for several reasons, including not knowing where exactly the city of interest is located, incorrect interpretation of weather symbols (e.g. dark clouds do not always imply rain), or excessive deterrence that could be caused by prominently displayed weather warning symbols or flashing lightnings. Therefore, this procedure should constitute a sound way to model how viewers of weather forecasts interpret the presented information. Moreover, Liu et al. (2014) argue that deliberately subjective measurements of weather could be helpful in understanding how individuals decide between different travel options.

In addition to the aforementioned manual evaluation procedure which is subjective by design, we also employ a computerized evaluation of the weather maps. The main idea behind this more objective procedure is the following: The darker the forecasted clouds around one's home region on the weather map, the less people will use their bike. This implies that the highest bike ridership would be observed when no clouds are forecasted, and that the lowest bike ridership would occur when very dark clouds are forecasted. To incorporate this notion into our regressions model, we use a digital image-processing method to prepare the static weather maps for a subsequent analysis of darkness values. Through digital image processing, we first recolor green land areas and blue sea areas on the maps in white[4] and then, the weather maps are converted to grayscale. The city labels are then removed from the map and the resulting empty areas are reconstructed using image interpolation.[5] This procedure is outlined for two different weather maps in Figure 3. In the next step, the resulting grayscale map without city labels is used to calculate the average darkness value (from 0 to 1) for the area that surrounds a given city. For example, this area is displayed through the red rectangles in Figures 3b and 3d for the city Hanover and spans roughly 7,500 km^2. In this example, the calculated values are 0.522 for Figure 3b, and 0.327 for Figure 3d. Then, the so-called *forecasted_darkness* values are linked with the analyzed city and, depending on whether the 8:00, 12:00 or 16:00 o'clock weather map was evaluated, with the corresponding hours of the observations as described above.

[4] This is done, because there is a clear sky in green or blue areas, which should indicate the best weather conditions for biking. The areas in green or blue, however, can have higher darkness values than areas where there are lighter clouds. Thus, an analysis of darkness values for unprocessed weather maps could indicate that areas with light clouds offer better biking conditions than areas with a clear sky. In order to solve this problem, the green and blue areas are recolored in white, which is the least dark color and therefore indicates the best weather conditions. It should also be noted that the shades of the clouds are not explicitly accounted for, meaning that they are also recolored if the shaded areas are too green (which is very often the case). From this follows also that the computerized evaluation calculates darkness values based on where the clouds are printed, not where their shadows might indicate their locations. This again reflects the main idea that cyclists would be less willing to use their bike if the respective region appears darker on the weather map. As clouds on the television weather forecast are moving objects, we further argue that it is reasonable to assume that the viewers would also locate the clouds exactly where they are printed, as the clouds themselves are much more visible and detectible than the clouds' rather transparent shadows.

[5] To do this, we employ the *inpaint* command from R's "imager" package. It uses a gaussian filter to replace the missing areas with a weighted average of neighboring regions.



(a) July 4th, 2017 – 8:00

(b) July 4th, 2017 – 8:00

(c) November 25th, 2017 – 16:00

(d) November 25th, 2017 – 16:00

Source: ARD / Hessischer Rundfunk.

Figure 3. : Original and processed weather maps for two sample days

Additional data sources

It is also important to account for official holidays, school holidays, and semester breaks. As the impact on bicycle counts could differ for each of these days, three separate dummy variables are included in the regression model.

The dummy variable *official_holidays* indicates if a day is an official holiday in the city's state, and the dummy variable *school_holidays* indicates if a day is a school holiday in the city's state. Moreover, the dummy variable *semester_break* indicates if a day lies in the semester break period of the largest university of the city where the bicycle counting station is located. If the bicycle counting station is located in a city or region with no university, the dummy variable always takes on a value of 0 for this bicycle counting station.

Preparation of the data

After linking the hourly count data with the corresponding weather information, weather forecast information, and the additional data sources outlined above, we restrict our sample to observations from 6:00 to 18:00 o'clock. This restriction on daytime counts is common in the literature, and in our case, also due to the information that we can derive from the historical weather forecast maps.

The dataset is scanned for implausible extreme values, which are subsequently excluded from the regression analysis (similar to Miranda-Moreno et al., 2013).

B. Descriptive statistics

Bicycle counting stations and their classification

Given that we use data from 188 bicycle counting stations in 37 cities, it is important to note that there are significant differences across cities and across counting stations. A detailed overview of descriptive statistics for the different cities in the sample can be found in Table 3 in Appendix A. This overview serves as an orientation on the general usage of bicycles within the various cities. It should be noted, however, that these bicycle counts depend on the actual location of the counting stations, and that they can therefore only be used as indicator of the average usage of bicycles within a city.

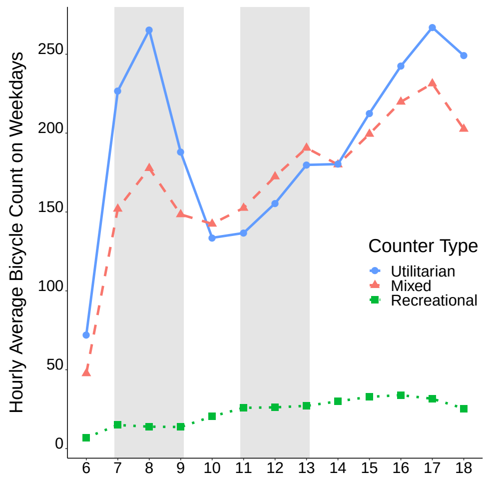
Besides differences at the city level, which can be seen in Table 3, there are also differences across counting stations. While bicycle traffic at some counting stations can be considered as mostly utilitarian, there are also stations in regions where people cycle mainly for recreational purposes. There are also stations where utilitarian and recreational bicycle traffic mix. Thus, we group the counting stations into three different types: utilitarian, mixed, and recreational. For this classification, which is inspired by Miranda-Moreno et al. (2013), we compare bicycle counts during morning peak hours on weekdays (07:00 – 09:00) to counts during noon hours on weekdays (11:00 – 13:00), as well as bicycle counts during weekdays to bicycle counts during weekends. To be classified as a utilitarian counting station, bicycle traffic at the considered station must satisfy two conditions: (U1) traffic during morning peak hours > traffic during noon hours, and (U2) traffic during weekdays > traffic during weekend days. To be classified as a recreational counting station, the following two conditions must be satisfied: (R1) traffic during morning peak hours < traffic during noon hours, and (R2) traffic during weekdays < traffic during weekend days. All bicycle counting stations that cannot be classified as either utilitarian or recreational, are subsequently assigned to the mixed group.

The average bicycle traffic profiles for utilitarian, mixed, and recreational counting stations can be found in Figure 4.

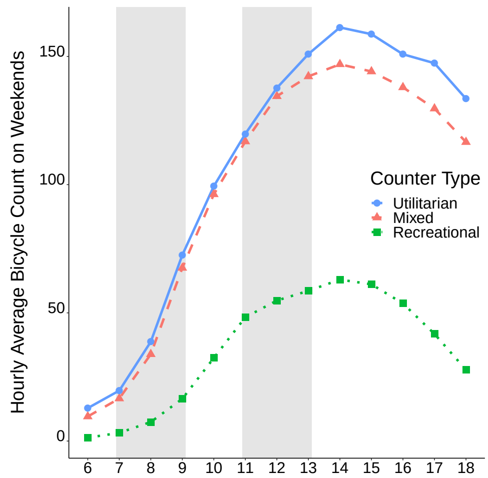
Here, the hourly profiles are displayed in Figure 4a for weekdays, and in Figure 4b for weekends. Daily profiles can be found in Figure 4c. Of the 188 bicycle counting stations, 122 can be classified as utilitarian, 34 as mixed, and 32 as recreational.

Weather variables

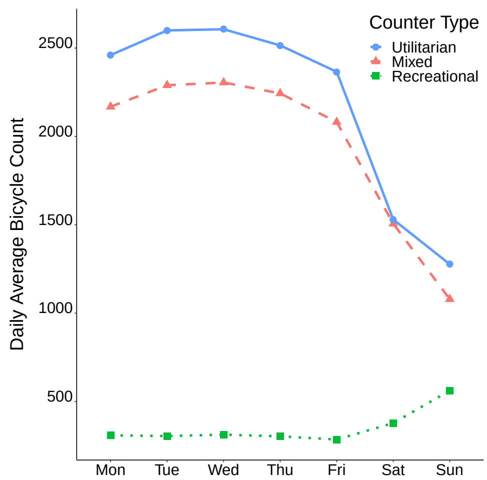
According to data from Beck et al. (2018), Germany has no dry seasons and a warm summer. While western and northern regions can be classified as temperate,



(a) Hourly profile for a weekday



(b) Hourly profile for a weekend day



(c) Daily profile

Figure 4. : Hourly and daily counter profiles

eastern and southern regions are rather cold. Following the Kppen-Geiger climate classification, the former regions thus can be classified as having a temperate oceanic climate (Cfb), while the latter regions can be classified as having a warm-summer humid continental climate (Dfb).

The average monthly air temperature, as well as the monthly precipitation for the whole sample period, can be seen in Figure 5. Generally, 2018 was a very hot year for Germany, with much less precipitation than usual.

Additional descriptive statistics of the actual weather variables and the weather forecast variables can be found in Table 4 in Appendix B.

IV. Methodology

Bicycle counts are a positively skewed count variable. In the literature, both log-linear and negative binomial regression models have been shown to provide a

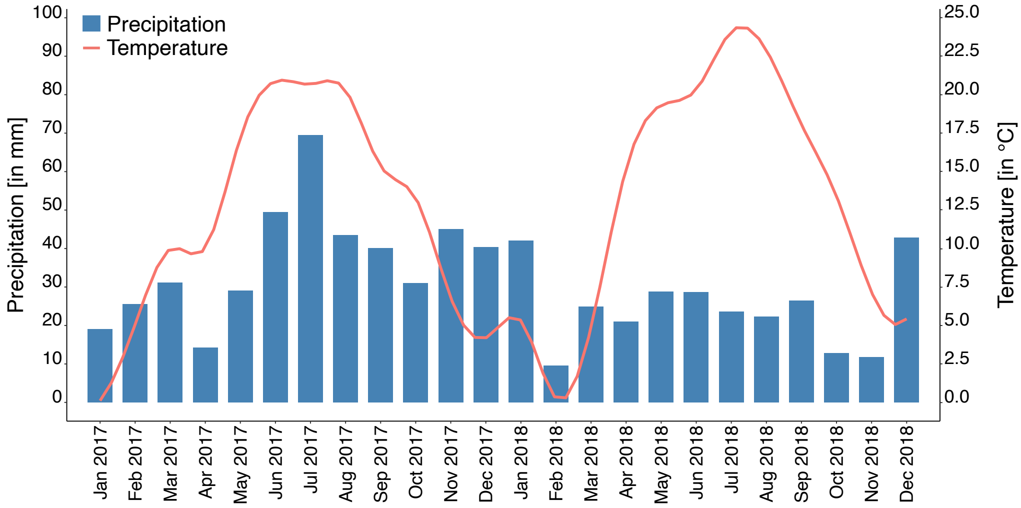


Figure 5. : Monthly weather profile

good fit for bicycle count data. They are used, for example, by Miranda-Moreno and Nosal (2011) or Nosal and Miranda-Moreno (2014). Moreover, Nordback (2012) compares the fit of different regression model types and shows that log-linear and negative binomial models provide the best fit for bicycle count data. Consequently, a log-linear regression model and a negative binomial regression model are estimated.

To analyze the different impacts of actual and forecasted weather variables, three different types of regression models are compared. The first only includes actual weather variables, and the second regression model only includes forecasted weather variables. The third regression model allows for a simultaneous inclusion of actual and forecasted weather variables.

For each type of regression model, different combinations of variables were tested in order to find the specification with the best fit for the regression model. The evaluation and comparison of different specifications was conducted using $\bar{R}^2$ or the AIC. As the third regression model features both actual and forecasted weather variables, we also ensure avoiding multicollinearity by evaluating the variance inflation factors (VIFs) for different specifications. As a criterion, we use VIF < 5. This is also the reason why, for example, only the actual and not the forecasted air temperature is included in the third regression model.

Not only linear, but also quadratic terms were tested in various specifications. However, only the variables *actual_air_temperature*, *forecasted_air_temperature*, as well as the precipitation level appeared to have a non-linear impact. Similar to Miranda-Moreno and Nosal (2011), we check whether one or more precipitation dummies would offer a better fit than the continuous precipitation variable. As it turns out, the best fit can be achieved by including six dummy variables indicating light drizzle (precipitation $< 0.5\,mm/h$), strong drizzle ($0.5\,mm/h \leq$ precipitation $< 1\,mm/h$), light rain ($1\,mm/h \leq$ precipitation $< 2\,mm/h$), moderate rain ($2\,mm/h \leq$ precipitation $< 5\,mm/h$), heavy rain ($5\,mm/h \leq$ precipitation $< 10\,mm/h$), and very heavy rain ($10\,mm/h \leq$ precipitation). This scale is adapted from the rain intensity scale presented by Tokay and Short (1996), which is representative, according to Dunkerley (2008). The minor change to the original scale provides a better fit with the actual precipitation data of Germany. The reference category is zero precipitation. Moreover, another dummy variable is used in the regression model to control for actual snowfall. It takes the value 1 if we have precipitation and sub-zero air temperatures, and otherwise 0.

We also include fixed effects for the bicycle counting station, hour of day, day of week, month of year, and year. Thereby, we account for all observable and unobservable effects that are constant for a given bicycle station, a given hour of the day, a given day of the week, a given month of the year, and for the year. It should be noted that these fixed effects, in general, have high explanatory power. Standard errors are clustered at the level of the bicycle counting stations, so that error terms from different bicycle counting stations are assumed to be independent.

Observations with missing values appear to be evenly distributed across all bicycle counting stations and are consequently excluded from the regressions (Miranda-Moreno and Nosal, 2011). Since we use the logarithm of *bicycle_counts* as our dependent variable in the log-linear regression model, we have to change observations with non-missing zeros to 1. For the whole sample, this amounts to 3.7 % of observations (1.4 % for utilitarian stations; 1.8 % for mixed stations; 14.5 % for recreational stations). The adjusted count variable will be denominated as $log(bicycle\_counts\_adj)$. For the negative binomial regression models, the dependent variable is not log-transformed, so that zero counts can be used in the regression.

For a better interpretation of the results, exact semi-elasticities for dummy and linear continuous variables are calculated as $100 \times [exp(\beta) - 1]$ (Wooldridge, 2012, p. 192). Semi-elasticities for quadratic terms are calculated as $100 \times [\beta + 2 \cdot \beta \cdot \bar{X}]$, with $\bar{X}$ as the mean of variable X (Wooldridge, 2012, p. 196 ff.).

V. Analysis

A. *Regression analysis with either actual or forecasted weather conditions*

In Section II.D, it was outlined that the existing literature offers no comprehensive analysis of the effects that various forecasted weather variables can have on bike ridership. Related to this, the results of a regression model using only actual weather variables, as well as the results of two regression models using only forecasted weather variables are outlined in Table 1.

The results of the log-linear regression that features only actual weather variables are presented in Regression (1) of Table 1.

The actual air temperature has a nonlinear effect and the semi-elasticity at the mean is 3.04. Bike ridership increases with rising air temperatures and peaks at around 29.5C. If the air temperature increases even further, bike ridership begins to decrease.

As expected, the rain dummy variables are negative and a higher precipitation intensity leads to stronger decreases in bicycle counts. For heavy and very heavy rain, however, the negative impact weakens.[6] If it rained in the last three hours,

[6] This might seem counterintuitive, but rainfall intensity is higher for short-lived rain events and a large share of the total precipitation in a rainfall event tends to fall in a small fraction of the event duration (Dunkerley, 2008). Since precipitation is measured hourly in our sample, it could thus be that in hours with heavy or very heavy rain, there are also longer periods without rain, compared to hours with moderate, but ongoing falling rain. Moreover, heavy rain only occurs in 0.12 % of the sample hours, and very heavy rain only in 0.04 %.

Table 1—: Separate effects of actual and forecasted weather variables (log-linear regression)

	Actual weather	Forecasted weather	
		Manual evaluation	Computerized evaluation
	(1)	(2)	(3)
actual_air_temperature	0.053*** (0.002)		
actual_air_temperature^2	−0.001*** (0.0001)		
actual_rain_light_drizzle	−0.109*** (0.006)		
actual_rain_strong_drizzle	−0.169*** (0.008)		
actual_rain_light_rain	−0.245*** (0.010)		
actual_rain_moderate_rain	−0.291*** (0.014)		
actual_rain_heavy_rain	−0.254*** (0.025)		
actual_rain_very_heavy_rain	−0.067** (0.033)		
actual_rain_in_last_3_hours	−0.277*** (0.010)		
actual_relative_humidity	−0.006*** (0.0005)		
actual_windspeed	−0.033*** (0.002)		
actual_cloudiness	−0.011*** (0.002)		
actual_snow	−0.034 (0.028)		
forecasted_air_temperature		0.048*** (0.002)	0.049*** (0.002)
forecasted_air_temperature^2		−0.0004*** (0.0001)	−0.0005*** (0.0001)
forecasted_light_clouds		−0.114*** (0.006)	
forecasted_dark_clouds		−0.209*** (0.010)	
forecasted_darkness			−0.538*** (0.020)
forecasted_rain		−0.382*** (0.015)	−0.125*** (0.006)
forecasted_thunderstorm		−0.365*** (0.020)	−0.185*** (0.013)
forecasted_snow		−0.314*** (0.013)	−0.040*** (0.013)
forecasted_warning		−0.131*** (0.006)	−0.136*** (0.007)
official_holidays	−0.783*** (0.045)	−0.751*** (0.046)	−0.769*** (0.046)
school_holidays	−0.192*** (0.010)	−0.196*** (0.010)	−0.195*** (0.010)
semester_break	−0.155*** (0.014)	−0.165*** (0.013)	−0.162*** (0.013)
Observations	1,576,348	1,599,738	1,599,738
Adjusted R^2	0.785	0.776	0.777
Residual Std. Error	0.780 (df = 1576114)	0.795 (df = 1599509)	0.793 (df = 1599510)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: Dependent variable: *log(bicycle_counts_adj)*. Fixed effects for the bicycle counting station, hour of day, day of week, month of year, and year are included at all times. Standard errors are clustered at the level of the bicycle counting stations.

bike ridership decreases by 24.2 %. Relative humidity, windspeed, and cloudiness all have negative impacts on hourly bike ridership. Actual snow, however, appears to have no significant impact in this regression model. Official holidays decrease bike counts by 54.3 %, school holidays have a negative impact of 17.5 %, and semester breaks a negative impact of 14.4 % on hourly bike ridership. It should be noted that the effects of holidays and semester breaks are similar for all three regressions in Table 1.

The results of the regressions with forecasted variables are presented in Regressions (2) and (3) of Table 1. Both regression specifications use variables derived from *Tagesschau* forecasts, with the difference that (2) only uses the manual evaluation of the weather forecasts (variables *forecasted_light_clouds* and *forecasted_dark_clouds*), whereas (3) uses the computerized evaluation of the forecasted cloud coverage (variable *forecasted_darkness*) as outlined in Section III.A. The forecasted air temperature has a positive, non-linear effect in both specifications, and the semi-elasticities at the mean are 3.45 and 3.46, respectively. In contrast to the actual air temperature, bike ridership increases monotonically over the relevant forecasted air temperature values. Since a region's forecasted day highest air temperature is constant over all hours of one day, it appears plausible that the curvature for the actual air temperature, which varies throughout one day, is higher than for the forecasted air temperature.

Also, the forecasted cloud coverage has a significant negative impact. This holds for the manual, as well as for the computerized evaluation of the *Tagesschau* weather maps. In Regression (2), the presence of light clouds in the region of the bicycle counting station decreases hourly bicycle ridership by 10.8 %, and dark clouds lead to a decrease of 18.9 %. For the computerized evaluation of the weather forecasts in Regression (3), the variable *forecasted_darkness* indicates the darkness of the area around the counting station on the weather forecast map. It takes the value of 0 for no cloud cover, and 1 for pitch-black clouds. Thus, the regression coefficient of -0.538 implies that if pitch-black clouds were forecasted, bike ridership would decrease by 41.6 % in comparison to a forecasted clear sky. A decrease from the $Q1$ value of *forecasted_darkness* to the $Q3$ value, that is from a darkness value of 0.00321 to 0.35224, would cause bike ridership to drop by 14.5 %. Therefore, it can be concluded that the forecasted cloud coverage significantly impacts on people's bike riding behaviour.

In addition to cloud coverage, forecasted rain, thunderstorms, and snow also have significant negative impacts on hourly bike ridership. These effects are stronger for Regression (2) than for Regression (3), which plausibly suggests that the variable *forecasted_darkness* already captures some of the rain, thunderstorm, and snow effects, as these three weather conditions are usually connected with darker cloud coverage on weather maps. If a weather warning symbol is displayed for the region of the counting station, hourly bike ridership decreases by roughly 12.5 %.

Additionally, a negative binomial regression model is used. In contrast to the poisson regression model, it does not assume that the mean and the variance are equal and thus it allows for overdispersion, which is present in the data according to a likelihood ratio test that compares the likelihood of the poisson model to the likelihood of the negative binomial model. The results of the negative binomial regressions (8) to (10), which are presented in Table 5 in Appendix C, are very similar and confirm the results of the log-linear regressions. One notable difference, however, is that actual snowfall has a significant negative impact in the negative binomial regression model.

B. Regression analysis with both actual and forecasted weather conditions

From Table 1, it becomes evident that both actual weather conditions and forecasted weather conditions significantly impact on hourly bike ridership. Thus, these variables are important when it comes to predicting hourly bicycle ridership. It is, however, not reasonable to assume that bike ridership is influenced by either only actual, or only forecasted weather conditions. More realistically, bicyclists decide to use their bike based on both forecasted, as well as actual weather conditions. This is reflected in the combined regression model presented as Regression (4) in Table 2. With the help of the combined regression model, we can also analyze the impact of wrong weather forecasts, which was outlined as the second research gap in Section II.D.

Table 2—: Weather effects by counter type (log-linear regression)

	Overall	by counter type		
		Utilitarian	Mixed	Recreational
	(4)	(5)	(6)	(7)
actual_air_temperature	0.055***	0.051***	0.049***	0.081***
	(0.002)	(0.002)	(0.004)	(0.005)
actual_air_temperature^2	−0.001***	−0.001***	−0.001***	−0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
actual_rain_light_drizzle	−0.158***	−0.136***	−0.177***	−0.194***
	(0.008)	(0.008)	(0.016)	(0.021)
actual_rain_strong_drizzle	−0.199***	−0.191***	−0.226***	−0.202***
	(0.009)	(0.010)	(0.020)	(0.027)
actual_rain_light_rain	−0.264***	−0.258***	−0.298***	−0.234***
	(0.012)	(0.014)	(0.022)	(0.029)
actual_rain_moderate_rain	−0.303***	−0.315***	−0.330***	−0.254***
	(0.015)	(0.015)	(0.032)	(0.032)
actual_rain_heavy_rain	−0.285***	−0.256***	−0.330***	−0.303***
	(0.025)	(0.027)	(0.047)	(0.054)
actual_rain_very_heavy_rain	−0.095***	−0.123***	−0.197**	−0.163*
	(0.033)	(0.035)	(0.097)	(0.084)
actual_rain_in_last_3_hours	−0.241***	−0.199***	−0.221***	−0.362***
	(0.009)	(0.007)	(0.016)	(0.021)
wrong_rain_forecast	−0.037***	−0.018***	−0.049***	−0.082***
	(0.005)	(0.005)	(0.008)	(0.010)
wrong_dry_forecast	0.109***	0.101***	0.110***	0.136***
	(0.008)	(0.010)	(0.018)	(0.013)
forecasted_thunderstorm	−0.034***	−0.011	−0.008	−0.098***
	(0.009)	(0.011)	(0.019)	(0.021)
forecasted_warning	−0.060***	−0.079***	−0.072***	−0.001
	(0.006)	(0.005)	(0.011)	(0.015)
actual_cloudiness	−0.005***	−0.003***	−0.0003	−0.016***
	(0.001)	(0.001)	(0.002)	(0.004)
forecasted_darkness	−0.268***	−0.236***	−0.223***	−0.468***
	(0.015)	(0.013)	(0.033)	(0.033)
actual_relative_humidity	−0.005***	−0.004***	−0.004***	−0.009***
	(0.0005)	(0.0004)	(0.001)	(0.001)
actual_windspeed	−0.028***	−0.022***	−0.025***	−0.044***
	(0.002)	(0.001)	(0.003)	(0.006)
official_holidays	−0.789***	−1.018***	−0.990***	0.271***
	(0.045)	(0.031)	(0.091)	(0.054)
school_holidays	−0.190***	−0.243***	−0.154***	−0.083***
	(0.010)	(0.011)	(0.015)	(0.015)
semester_break	−0.156***	−0.131***	−0.169***	0.054
	(0.014)	(0.013)	(0.038)	(0.055)
Observations	1,576,348	1,029,717	272,669	273,962
Adjusted R^2	0.786	0.721	0.829	0.739
Residual Std. Error	0.778 (df = 1576110)	0.711 (df = 1029545)	0.664 (df = 272585)	0.813 (df = 273880)

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Note: Dependent variable: *log(bicycle_counts_adj)*. Fixed effects for the bicycle counting station, hour of day, day of week, month of year, and year are included at all times. Standard errors are clustered at the level of the bicycle counting stations.

The actual air temperature has, similar to Regression (1), a positive, but non-linear effect on hourly bike ridership. Due to high VIF values, the forecasted air temperature is excluded from the regression in order to avoid multicollinearity issues. The negative effects for different rain categories are comparable to Regression (1), yet slightly higher. If it rained in the previous three hours, bicycle traffic decreases by 21.4 %.

Forecasted rain is not included directly in the regression model, but we test whether wrong rain forecasts impact actual bike ridership. The regression coefficient for the variable *wrong_rain_forecast* gives the percentage change in bike ridership if rain is forecasted, but it then does not rain in reality. The regression coefficient for the variable *wrong_dry_forecast* gives the percentage change in bike ridership if no rain is forecasted, but it actually does rain. The regression results imply that falsely forecasted rain can reduce bike ridership by 3.6 %. On the other hand, bike ridership in rainy hours will be 11.5 % higher if no rain is forecasted, compared to the case if rain was correctly forecasted. These findings suggest that bicyclists base their travel decisions not only on the actual weather conditions, but also on information from weather forecasts – and, more importantly, some bicyclists appear to stick to these decisions even if the actual weather conditions are different from the forecasts.

Similar to the regression models with only forecasted weather conditions, the effects of forecasted thunderstorms and forecasted weather warnings are also negative and significant in this combined regression model.

Also, Regression (4) shows that the actual cloudiness, measured on a scale from 0 (clear sky) to 8 (8/8 of the sky are covered by clouds), decreases bike ridership by 0.5 % per eighth of the sky that is covered by clouds. Thus, bike ridership under a clear sky would be 4 % higher than under a fully covered sky. In addition to the rather moderate effect of actual cloudiness, the darkness of forecasted clouds on the weather map can decrease bike ridership by 23.5 % if the forecasted weather map shows a pitch-black cloud cover instead of a clear sky. An increase from the $Q1$ value of *forecasted_darkness* to the $Q3$ value would cause bike ridership to drop by 8.2 %.

The effects of actual relative humidity and actual windspeed are similar to those in Regression (1). The same holds for the effects of official holidays, school holidays, and semester breaks.

The results of the corresponding negative binomial regression, which is presented as Regression (11) in Table 6 in Appendix C, are mostly similar to the log-linear regression. Two notable differences are, however, that very heavy rain now has no significant negative impact on hourly bike ridership and that the positive effect of falsely predicted dry weather is slightly smaller.

C. *Regression analysis for different counter types*

To test the robustness of Regression (4), we estimate if weather effects are constant over different counter types and, thus, provide additional insights into the third research gap of Section II.D. As outlined in Section III.B, the 188 bicycle counting stations of our sample can be classified as either utilitarian, recreational, or mixed. Figure 4 showed that these three types of counting stations differ with respect to their hourly and daily profiles. Thus, it can be expected that the counting stations also differ with respect to the estimated weather effects. Regression (5) shows the estimated coefficients for utilitarian counters, (6) for mixed counters, and (7) for recreational counters.

Here, it is interesting to observe that hourly bicycle ridership at recreational counting stations is more strongly affected by air temperature increases than utilitarian and mixed bike traffic (semi-elasticity of 5.63 % at the mean, compared to 2.88 % and 2.84 %). Moreover, bicycle ridership at recreational counting stations decreases by 30.4 % if it rained in the previous three hours, but only 18.0 % at utilitarian stations. The effect of falsely forecasted rain is only −1.8 % at utilitarian stations, but −4.8 % at mixed stations and even −7.9 % at recreational stations. Moreover, the negative impacts of forecasted thunderstorms, actual cloudiness, and of the darkness of the respective regions on the weather forecast maps are much more pronounced for recreational stations than for utilitarian and mixed ones. Thus, the results of this sensitivity analysis suggest that recreational bike traffic is more sensitive to actual and forecasted weather, especially to air temperatures, rain in previous hours, falsely forecasted rain, and the darkness of the weather forecast maps.

Besides weather effects, the three types of counting stations also differ with respect to the impact of holidays and semester breaks. While utilitarian and mixed bicycle traffic decrease substantially on official holidays (−63.9 % and −62.8 %), hourly bike ridership at recreational counting stations increases by 31.1 %. Moreover, the negative impact of school holidays is stronger for utilitarian than for recreational counting stations, and while semester breaks decrease utilitarian and mixed bicycle traffic, they seem to have no significant impact on recreational stations. Here, it should be noted that some of the recreational bicycle stations are located in regions where there are no large universities, so that the *semester_break* dummy is constantly zero for these stations. All in all, the results of these sensitivity analyses for the three different types of bicycle counting stations are mostly plausible and intuitive. Consequently, they validate the overall regression setup and also the differentiation between the three types of counting stations.

Again, the results of the negative binomial sensitivity analysis are mostly similar. These results are presented as Regressions (12) to (14) of Table 6 in Appendix C.

D. *Lagging and leading effects of weather forecasts*

As of now, we have analyzed the impact that weather forecasts have on the number of bicyclists in the hours for which the weather conditions were forecasted. It might be possible, however, that weather forecasts also have a lagging and/or a leading effect on bike ridership. For example, if rain was forecasted for the morning hours, commuters might decide to not use their bike for their morning commute. Subsequently, they would not be able to use their bike on their way back from work, thus also reducing bicycle ridership in afternoon hours. A more extensive analysis of these effects is presented in Table 7 in Appendix D. These results help to provide answers to the fourth research gap of Section II.D.

In Regression (15), we test if the impact of rain forecasts on bike ridership is restricted to the hours for which the rain is forecasted, or if, for example, forecasted morning rain has an effect on all-day bike ridership. To test this, the variable *forecasted_morning_rain* takes on the value 1 if rain is forecasted for only the morning hours of that day, and otherwise 0. Its regression coefficient implies that if rain is forecasted for morning hours only, the bike ridership for the whole day will be 9.2 % lower. Also, rain forecasts for midday hours only have negative effects on bike ridership over the whole day, but to a lesser degree (−5.2 %). If rain is forecasted for only the afternoon hours, then bike ridership for the whole day will not significantly decrease. Thus, it can be concluded that rain forecasts for morning and midday hours can influence bike ridership over the whole day, whereas rain forecasts for afternoon hours have no significant impact on all-day bike ridership.

This notion is further analyzed with Regression (16), which estimates the effects that rain forecasts can have on the hours for which the rain is forecasted, and also for other hours of the day that were predicted to be rain-free. Here, the variable *forecasted_morning_rain_on_morning_traffic* captures the effect that a rain forecast for the morning hours has on bike ridership in the morning hours. It takes on the value 1 for morning hour observations for which rain is forecasted, and otherwise 0. The variable *forecasted_morning_rain_on_midday_traffic* captures the effect that morning rain forecasts have on bike ridership in midday hours for which no rain was forecasted. It takes on the value 1 for midday hour observations of days for which rain is forecasted for the morning hours, but not for the midday hours; otherwise, it takes on the value 0. The other direct and time-shifted weather forecast effect variables are created similarly. The results indicate that if rain is forecasted for the morning hours, bike ridership in morning hours is 25.6 % lower compared to a rain-free weather forecast. The results also show that rain forecasts for the morning hours reduce not only bike ridership during this time, but also during midday and afternoon hours that were predicted to be rain-free. This lagging effect is strong and lies around −14.7 % for midday hours and −12.6 % for afternoon hours. If rain is forecasted for the midday hours, it reduces bike ridership in midday hours by 33.5 %. Furthermore, midday rain forecasts also reduce bike ridership in morning hours that were predicted to be rain-free by 1.5 %, and they reduce bike ridership in afternoon hours that were predicted to be rain-free by 9.6 %. Rain forecasts for afternoon hours have a strong effect on bike ridership in afternoon hours (−30.4 %), but they do not significantly affect bike ridership in morning hours that were predicted to be rain-free. However, bike ridership in midday hours that were predicted to be rain-free is slightly reduced by afternoon rain forecasts (−1.4 %). The results of Regression (16) thus show that we have a pronounced lagging effect of rain forecasts. Especially rain forecasts for the morning hours can significantly reduce bike ridership in subsequent hours. One explanation could be that if commuters expect to get wet on their morning commute, they avoid using their bike for this trip and subsequently cannot use it on their home commute or on later trips that start from their workplace. For actual rain, such a lagging effect was already found in other studies (e.g. Miranda-Moreno and Nosal, 2011) and also confirmed in this study by the variable *actual_rain_in_last_3_hours*. However, the aforementioned results underline that the lagging effect is not only existent for actual rain, but also for forecasted rain. Besides this lagging effect, we can also see less pronounced leading effects of midday rain forecasts on morning bike ridership and of afternoon rain forecasts on midday bike ridership.

In Regression (17), we test if all-day rain forecasts reduce bike ridership stronger than rain forecasts for just a few hours. For morning hours, the variable *forecasted_morning_rain_only_on_morning_traffic* takes on the value 1 if morning rain is forecasted, and no midday or afternoon rain is forecasted as well. The variable *forecasted_morning_rain_plus_one_on_morning_traffic* takes on the value 1 if morning rain as well as either midday or afternoon rain for the same day is forecasted. The variable *forecasted_all_day_rain_on_morning_traffic* takes on the value 1 if rain is forecasted for the whole day, that is for morning, midday, and afternoon hours. The results show that midday and afternoon traffic is reduced if rain is forecasted for the respective hours of the day (−26.3 % for midday hours; −20.9 % for afternoon hours), but it is even stronger reduced if rain is also forecasted for some other hours of the day (−29.1 %; −28.5 %). If rain is forecasted for the whole day, the reduction in bike ridership is the most pronounced (−35.0 %; −33.6 %). For morning hours, however, this effect is surprisingly reversed. If rain is only forecasted for the morning hours, then bike ridership decreases by 26.8 %, but it decreases by only 24 % if rain is forecasted for some additional hours of the day. If rain is forecasted for the whole day, bike ridership in morning hours is only reduced by 20.9 %. One potential explanation for this unexpected decrease in effect magnitude might be that people who have to use their bikes for certain midday or afternoon trips could reschedule these trips to the morning hours, as they would expect to get wet anyway if rain was also forecasted for later parts of the day. To verify this potential explanation, however, further research on the decision-making process of bicyclists is needed.

Similar results can be obtained when using a negative binomial regression model.[7] It should also be noted that this subsection focuses on leading and lagging effects of forecasted rain. For other forecasted weather conditions, the leading and lagging effects on bike ridership could be less clear due to the high discomfort that rain brings for bicyclists and its strong deterrence of cycling. A comprehensive analysis of the leading and lagging effects of other forecasted weather conditions, however, remains an interesting topic for further research.

E. Sensitivity analyses

Additional sensitivity analyses show that the main results are not significantly different if we look separately at the years 2017 and 2018. There are, however, differences if we look at summer and winter months separately. The results of this sensitivity analysis show that the effects of wrong weather forecasts, as estimated through the variables *wrong_rain_forecast* and *wrong_dry_forecast*, are stronger in summer months than in winter months. This finding indicates that in winter months, a higher share of the potential winter bicyclists use their bike regardless of actual and forecasted weather conditions, whereas in summer months, a larger share of the potential summer bicyclists are only willing to use their bike if weather conditions are generally more favorable. As good weather conditions are important to these people, they tend to check weather forecasts more regularly and are thus more sensitive to these forecasts.[8]

VI. Discussion and conclusions

A. Discussion of results and policy implications

First, we estimate the impact of several weather forecast variables on bike ridership. The forecasted daily high air temperature has a positive effect on bike ridership, whereas forecasted clouds, rain, snow, thunderstorms, or weather warnings reduce bike usage. The effects of temperature and rain are in line with Kraemer et al. (2015), whereas effects for other weather forecast variables have

[7] The detailed results of this negative binomial regression model are available upon request.
[8] Complete regression tables for these sensitivity analyses are available upon request.

not yet been studied. It is interesting to see that the two regression models with only forecasted weather variables have an adjusted R^2 that is slightly lower, yet still comparable to the one of the regression model with only actual weather variables. This underscores that weather forecast variables can be used as valuable predictors for bike ridership. Moreover, this finding is especially important for traffic planners, who should account for actual weather conditions as well as for forecasted conditions. Both are important determinants of bike ridership and, consequently, impact on the whole traffic system.

Second, bicycle usage is influenced by wrong weather forecasts, that is if forecasted and actual weather conditions diverge. This underlines that weather forecasts can significantly influence the decision to use the bike or not. In accordance with this, the adjusted R^2 of the combined regression models is slightly higher than for the regression models using either only actual or only forecasted weather variables. The fact that bike ridership is lower in rain-free hours for which rain was forecasted compared to rain-free hours that were predicted to be rain-free, indicates that the decision to use the bike or not can lead to a "lock-in" effect. If a potential bicyclists watches the evening weather forecast and, based on the poor outlook, commits to driving with a shared car or with an arranged carpool on the next day, it would be difficult to still change this decision the next morning. In addition to this, the fact that next-day weather forecasts influence actual bike ridership – irrespective of whether they are right or wrong – should be considered when attempting to promote cycling. Therefore, it should be emphasized more that people should, if possible, not make their travel decisions based on next-day weather forecast maps, but rather rely on information from more timely forecast types such as rainfall radars. Otherwise, driving pleasure could easily turn to discomfort and frustration if the weather conditions are different than expected, thereby inducing potential bicyclists to switch to more weather-sheltered and reliable transport modes in the long-term.

Third, it is shown that the effects of actual and forecasted weather variables are stronger for recreational counting stations than for utilitarian ones. Utilitarian trips like daily commutes are often less flexible, implying that they cannot easily be cancelled or rescheduled if the weather is bad. Recreational trips, on the other hand, are more flexible and can be cancelled or postponed to a time when the weather is likely to be better. Thus, utilitarian trips are less elastic than recreational trips with regard to adverse weather conditions. This result is important for traffic planners, who need to account for these effect differences in order to allow for a smooth traffic system. Ignoring these effect differences – or the impact of weather forecasts in general – could lead to wrong predictions of actual traffic flows and thus reduce the efficiency of the traffic system. Moreover, if policymakers want to promote cycling, they need to tailor their promotional measurements to the type of traffic that is prevalent in their region. Otherwise, promotional measurements could fail to achieve their goals, as the impact of such measurements could differ for utilitarian and recreational traffic.

Fourth, it is shown that rain forecasts have various lagging and leading effects on bike ridership. The strongest effect can be observed if rain is forecasted for the morning hours, because this significantly reduces bike ridership in midday and afternoon hours, even if they are predicted to be rain-free. It is understandable that forecasted morning hour rain has the strongest lagging effect since the modal choice for trips that are subsequent to the morning commute often depends on whether a bike is available at the time of that trip, or if it is still at home. In order to mitigate this lagging effect of forecasted morning rain, it is of special importance to reduce the discomfort of morning commutes in the rain. One way to pursue this could be through faster changes from red to green-light phases at intersections if it rains and cyclists are approaching the intersection (Jorna and Zoer, 2012). Moreover, employers could provide showers and locker rooms to allow employed bicyclists to change their clothes after getting wet on the way to work. It is shown that the provision of showers can increase the attractiveness of cycling (Hunt and Abraham, 2007). In combination with similar measures, the outlook on morning commutes in the rain could become less discomforting for cyclists and thus help to increase bike ridership as well as to decrease car traffic in midday and afternoon hours.

In general, the results of the above analyses underline that not only actual rain, but even the mere prospect of getting wet can lead people to not use their bike. These results thus underline that if policymakers want to further promote cycling, they should try to alleviate the discomfort caused by rain. This could be achieved by reducing the exposure to rainfall. In addition to faster green-light phases, cycleways could be planned in such a way that adjoining trees with larger leaves could serve as a natural rain protection for cyclists. Tree canopies can have different rainfall interception rates, depending on the leaf sizes of the respective tree species (Yang et al., 2019).[9] Another example would be to plan cycleways as lanes that are separated from road traffic, so that the risk of splash water from cars is reduced. Moreover, introducing programs to make bicycle rain gear available at lower prices could increase the rain resilience of cyclists. These and other measures could help to make cycling in the rain less unpleasant, and thus help to mitigate the negative effects that rainy weather and rain forecasts have on bike ridership.

B. Discussion of the used methods and further research directions

One potential drawback of the study is that, due to data availability, we had to derive weather forecast data from historical weather maps that were broadcasted by Germany's most-watched television news program *Die Tagesschau*. Thus, we evaluated only one of many potential sources from which people could access weather forecast information. Due to the large reach of the *Tagesschau*, we nevertheless believe that our weather forecast variables are a good approximation of the general weather forecast information that people can access when making their travel decisions. This notion is also supported by Cools and Creemers (2013), who show that behavioral travel adaptions are independent of the manner in which weather information is acquired (media source, exposure time, perceived reliability).

Also, weather forecast maps are evaluated manually and might therefore be prone to subjective bias. In Regression (3), however, the manual evaluation of cloud coverage is substituted by a computerized evaluation using a digital image-processing method. The results confirm that darker clouds have a negative effect

[9] Although tree canopies can reduce exposure to immediate rainfall, they might also lead to cyclists being exposed to fewer, but larger raindrops after the actual rainfall period. Providing cyclists with a cycle lane network that allows them to choose between routes that are sheltered by tree canopies and routes that are not sheltered by tree canopies might thus be ideal with respect to cycling comfort on rainy days. Feasibility or efficiency of such a cycle lane network might, however, not always be given. Thus, additional research on these aspects and the effects of tree canopies on cyclists provide an interesting future research direction.

on bike ridership – independent of whether cloud coverage was evaluated manually or computerized.

As outlined earlier, the adjusted R^2 of the combined regression models is slightly higher than for the regression model types using either only actual or only forecasted weather variables. It should, however, be noted that the different regression model types also account for slightly different weather conditions. For example, the actual weather models do not account for weather warnings, whereas the forecasted weather models do not account for windspeed or relative humidty. These differences in variable selection are mainly due to data availability, and it would be interesting to analyze how the different regression model types would compare to each other if the exact same weather conditions were accounted for in the actual, the forecasted, as well as in the combined regression models.

By using automated bicycle count data, we can only estimate effects on the number of bikes that drive by a certain point, which usually is the location of the counting station. While our count data are a reliable data source and provide a reasonable approximation of actual bike ridership, they cannot be used to estimate how the forecasted weather impacts on certain characteristics of a trip. For example, forecasted rain or wind might not only lead to a decrease in bike ridership, but also to a change in trip routes and trip lengths if bicyclists could thereby reduce the exposure to adverse weather conditions. Using GPS data, or data from surveys and travel logs could provide further insights into these effects.

Another important research direction could be to further study individuals decision-making behavior based on the weather forecast data that they consume. It is necessary to better understand how individuals perceive different weather forecast types and how these perceptions lead to changes in their cycling trips. Especially with regard to lagging and leading effects of weather forecasts, it would be useful to extend our results, which are based on count data at a non-individual level, by an analysis of individual cycling behavior, for example through the use of GPS or travel log data. It would also be interesting to study how the subjective perception of the weather is not only influenced by the actual weather conditions, but also by (wrong) weather forecasts. Rain that was not forecasted could, for example, be perceived as more discomforting than forecasted rain, because cyclists are surprised by the bad weather and thus not able to protect themselves adequately.

Considering that rain often causes bicycle trips to be substituted by public transport or by car (Sabir, 2011) and that bike ridership will decrease if rain is forecasted, it could reasonably be assumed that rain forecasts cause comparable substitution effects. Such effects would, of course, depend on the quality of public transport and of the transport infrastructure. Nevertheless, it would be interesting for future research to estimate the substitution effects of weather forecasts and the ensuing environmental impacts.

C. Conclusions

This study uses hourly bicycle counts from 188 automated bicycle counting stations in 37 cities of Germany. For 2017 and 2018, these counts are connected to both actual and forecasted weather data.

The results show that both actual and forecasted weather conditions can be used as valuable predictors of hourly bike ridership. The novel results for forecasted weather conditions show that the forecasted air temperature has a positive effect on hourly bike ridership. Bicycle usage drops if clouds, rain, snow, thunderstorms, or weather warnings are forecasted. Also, the darkness of the forecasted clouds on weather forecast maps can reduce next-day bike ridership.

Bicycle usage is also influenced by wrong weather forecasts, that is if forecasted and actual weather conditions diverge. The results indicate that rain forecasts can reduce bike ridership in rain-free hours by 3.6 % and that forecasts of rain-free weather can increase bike ridership in rainy hours by 11.5 %.

It is shown that the effects of actual and forecasted weather variables on hourly bike ridership vary between utilitarian, recreational, and mixed counting stations. In general, they are stronger for recreational counting stations than for utilitarian ones.

Moreover, various lagging and leading effects of weather forecasts are outlined. Morning rain forecasts, for example, have a lagging negative effect on bike ridership in midday and afternoon hours, even if these hours were predicted to be rain-free.

Acknowledgements

The author would like to thank three anonymous referees, Kathrin Goldmann, and Gernot Sieg for their helpful comments; Kathrin Goldmann, Sebastian Specht, Alina Krämer, Carina Gugel, and Till Kösters for their help in collecting and preparing the data; Brian Bloch for his editing services; Peter Unterberg and Michael Magnus from Ecocounter for their helpful insights with regard to Ecocounter's automated bicycle counting stations; and the bike managers of the cities for letting us use the data from their bicycle counting stations.

Appendices

A. Descriptive statistics for bicycle counts in each city of the sample

Table 3—: Descriptive statistics of hourly bicycle counts by city

City	Observations	Mean	Min	Q1	Q2	Q3	Max
Augsburg	20722	120.2	0	55	103	173	510
Berlin	156640	195.9	0	55	123	262	1994
Bochum	4708	10.2	0	5	9	14	52
Bonn	141741	98.5	0	25	60	135	1059
Bremen	85090	222.1	0	104	186	305	1113
Cologne	103782	171.3	0	59	138	248	1237
Dsseldorf	100499	74.0	0	22	49	95	1200
Erlangen	12818	328.9	0	171	322	470	1478
Essen	5902	82.0	0	33	58	109	565
Freiburg	37831	328.2	0	116	213	499	1904
Gttingen	47307	169.0	0	74	147	235	1022
Hamburg	9464	376.9	2	158	306	513	1938
Hannover	82601	191.8	0	96	171	263	1147
Heidelberg	43316	240.7	0	103	206	338	1776
Heilbronn	18836	56.8	0	11	23	78	493
Herzogenaurach	5226	40.2	0	17	33	57	182
Jena	3874	81.1	0	24	56	112	486
Kiel	20982	47.2	0	20	38	65	398
Kirchheim unter Teck	9464	71.6	0	37	68	101	939
Leipzig	4732	239.3	0	114	224	341	758
Lrrach	10343	39.6	0	21	36	51	304
Ludwigsburg	4862	150.9	0	67	118	189	1065
Mannheim	28392	188.8	0	59	150	292	955
Munich	56784	110.6	0	23	68	153	1521
Mnster	73794	451.0	0	174	395	652	2919
Nationalpark Mritz	9464	12.5	0	0	1	15	171
Nuremberg	9464	128.1	1	69	118	174	462
Oberhausen	4940	31.3	0	12	23	44	499
Oldenburg	72373	130.3	0	40	87	156	1262
Rhein-Erft-Kreis	88678	25.8	0	4	11	33	426
Rhein-Kreis Neuss	46754	19.4	0	2	8	22	475
Rhein-Sieg-Kreis	94639	21.0	0	3	9	22	1151
Rostock	89726	58.0	0	11	35	76	1088
Sauerland Radwelt	38220	8.6	0	0	3	11	200
Stuttgart	9464	50.2	0	19	40	68	271
Tbingen	27378	218.6	0	45	149	349	1238
Wrzburg	18928	30.8	0	12	24	44	162
Overall	1599738	140.9	0	21	74	190	2919

B. Descriptive statistics for weather variables

Table 4—: Descriptive statistics for actual and forecasted weather variables

Variable	Observations	Mean	Min	Q1	Q2	Q3	Max
actual_cloudiness	1607600	5.773	0.0	4.0	7.0	8.0	8.0
actual_precipitation	1622946	0.077	0.0	0.0	0.0	0.0	48.8
actual_rain_light_drizzle	1626784	0.058		*(dummy variable)*			
actual_rain_strong_drizzle	1626784	0.020		*(dummy variable)*			
actual_rain_slight_rain	1626784	0.014		*(dummy variable)*			
actual_rain_moderate_rain	1626784	0.007		*(dummy variable)*			
actual_rain_strong_rain	1626784	0.0012		*(dummy variable)*			
actual_rain_heavy_rain	1626784	0.0004		*(dummy variable)*			
actual_relative_humidity	1621858	69.257	8.0	55.0	72.0	85.0	100.0
actual_snow	1626784	0.003		*(dummy variable)*			
actual_air_temperature	1622529	12.577	-16.6	5.9	12.4	19.1	37.3
actual_windspeed	1621166	3.775	0.0	2.2	3.4	4.9	22.4
forecasted_dark_clouds	1626784	0.162		*(dummy variable)*			
forecasted_darkness	1626784	0.210	0.0	0.0	0.1	0.4	0.9
forecasted_light_clouds	1626784	0.260		*(dummy variable)*			
forecasted_rain	1626784	0.163		*(dummy variable)*			
forecasted_snow	1626784	0.014		*(dummy variable)*			
forecasted_air_temperature	1626784	15.604	-9.0	8.0	16.0	23.0	38.0
forecasted_thunderstorm	1626784	0.018		*(dummy variable)*			
forecasted_warning	1626784	0.065		*(dummy variable)*			
official_holidays	1626784	0.029		*(dummy variable)*			
school_holidays	1626784	0.239		*(dummy variable)*			
semester_break	1626784	0.371		*(dummy variable)*			
wrong_rain_forecast	1626784	0.124		*(dummy variable)*			
wrong_dry_forecast	1626784	0.045		*(dummy variable)*			

Note: Descriptive statistics for the dummy variables that are used to estimate the lagging and leading effects of rain forecasts in Section V.D are available upon requests.

C. Negative binomial regressions

Table 5—: Separate effects of actual and forecasted weather variables (negative binomial regression)

	Actual weather	Forecasted weather	
		Manual evaluation	Computerized evaluation
	(8)	(9)	(10)
actual_air_temperature	0.046***		
	(0.002)		
actual_air_temperature^2	−0.0007***		
	(0.00006)		
actual_rain_light_drizzle	−0.098***		
	(0.004)		
actual_rain_strong_drizzle	−0.142***		
	(0.006)		
actual_rain_slight_rain	−0.214***		
	(0.008)		
actual_rain_moderate_rain	−0.253***		
	(0.013)		
actual_rain_strong_rain	−0.207***		
	(0.028)		
actual_rain_heavy_rain	0.052*		
	(0.029)		
actual_rain_in_last_3_hours	−0.259***		
	(0.009)		
actual_relative_humidity	−0.006***		
	(0.0005)		
actual_windspeed	−0.031***		
	(0.001)		
actual_cloudiness	−0.012***		
	(0.002)		
actual_snow	−0.145***		
	(0.022)		
forecasted_air_temperature		0.044***	0.045***
		(0.002)	(0.002)
forecasted_air_temperature^2		−0.0004***	−0.0004***
		(0.00004)	(0.00004)
forecasted_light_clouds		−0.117***	
		(0.005)	
forecasted_dark_clouds		−0.207***	
		(0.010)	
forecasted_darkness			−0.496***
			(0.018)
forecasted_rain		−0.346***	−0.109***
		(0.014)	(0.005)
forecasted_thunderstorm		−0.323***	−0.166***
		(0.017)	(0.011)
forecasted_snow		−0.325***	−0.066***
		(0.013)	(0.010)
forecasted_warning		−0.105***	−0.110***
		(0.005)	(0.005)
official_holidays	−0.516***	−0.467***	−0.483***
	(0.052)	(0.054)	(0.054)
school_holidays	−0.195***	−0.197***	−0.197***
	(0.009)	(0.008)	(0.008)
semester_break	−0.152***	−0.164***	−0.162***
	(0.012)	(0.012)	(0.012)
Observations	1,576,348	1,599,738	1,599,738
Squared Correlation	0.759	0.759	0.762

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: Dependent variable: *bicycle_counts*. Fixed effects for the bicycle counting station, hour of day, day of week, month of year, and year are included at all times. Standard errors are clustered at the level of the bicycle counting stations.

Table 6—: Weather effects by counter type (negative binomial regression)

	Overall	by counter type		
		Utilitarian	Mixed	Recreational
	(11)	(12)	(13)	(14)
actual_air_temperature	0.049***	0.045***	0.042***	0.113***
	(0.002)	(0.002)	(0.004)	(0.008)
actual_air_temperature^2	−0.0007***	−0.0007***	−0.0006***	−0.002***
	(0.00006)	(0.00004)	(0.00009)	(0.0002)
actual_rain_light_drizzle	−0.138***	−0.127***	−0.146***	−0.278***
	(0.006)	(0.005)	(0.014)	(0.023)
actual_rain_strong_drizzle	−0.164***	−0.166***	−0.171***	−0.296***
	(0.007)	(0.007)	(0.016)	(0.026)
actual_rain_light_rain	−0.231***	−0.225***	−0.264***	−0.322***
	(0.010)	(0.010)	(0.021)	(0.030)
actual_rain_moderate_rain	−0.267***	−0.269***	−0.282***	−0.329***
	(0.014)	(0.013)	(0.030)	(0.038)
actual_rain_heavy_rain	−0.242***	−0.213***	−0.268***	−0.345***
	(0.026)	(0.025)	(0.041)	(0.056)
actual_rain_very_heavy_rain	0.018	−0.010	−0.104*	−0.067
	(0.029)	(0.024)	(0.061)	(0.101)
actual_rain_in_last_3_hours	−0.225***	−0.190***	−0.203***	−0.446***
	(0.008)	(0.006)	(0.015)	(0.019)
wrong_rain_forecast	−0.035***	−0.024***	−0.042***	−0.077***
	(0.004)	(0.004)	(0.008)	(0.013)
wrong_dry_forecast	0.084***	0.087***	0.080***	0.149***
	(0.007)	(0.007)	(0.015)	(0.016)
forecasted_thunderstorm	−0.040***	−0.004	−0.020	−0.050**
	(0.008)	(0.007)	(0.016)	(0.020)
forecasted_warning	−0.053***	−0.071***	−0.060***	−0.069***
	(0.005)	(0.005)	(0.011)	(0.014)
actual_cloudiness	−0.006***	−0.004***	−0.002	−0.019***
	(0.001)	(0.001)	(0.002)	(0.004)
forecasted_darkness	−0.262***	−0.231***	−0.215***	−0.584***
	(0.013)	(0.010)	(0.027)	(0.031)
actual_relative_humidity	−0.004***	−0.004***	−0.003***	−0.006***
	(0.0004)	(0.0003)	(0.001)	(0.001)
actual_windspeed	−0.026***	−0.021***	−0.022***	−0.047***
	(0.001)	(0.001)	(0.002)	(0.005)
official_holidays	−0.517***	−0.769***	−0.759***	0.488***
	(0.052)	(0.037)	(0.107)	(0.056)
school_holidays	−0.193***	−0.228***	−0.152***	−0.094***
	(0.009)	(0.009)	(0.013)	(0.013)
semester_break	−0.154***	−0.121***	−0.175***	0.072
	(0.012)	(0.010)	(0.034)	(0.059)
Observations	1,576,348	1,029,717	272,669	273,962
Squared Correlation	0.762	0.786	0.810	0.749

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: Dependent variable: *bicycle_counts*. Fixed effects for the bicycle counting station, hour of day, day of week, month of year, and year are included at all times. Standard errors are clustered at the level of the bicycle counting stations.

D. Lagging and leading effects of weather forecasts

Table 7—: Lagging and leading weather forecast effects (log-linear regression)

	Rain forecast effects for whole day	Rain forecast effects for different daytimes	
		Single Effect	Combined Effects
	(15)	(16)	(17)
actual_air_temperature	0.054***	0.056***	0.054***
	(0.002)	(0.002)	(0.002)
actual_air_temperature^2	−0.001***	−0.001***	−0.001***
	(0.0001)	(0.0001)	(0.0001)
forecasted_thunderstorm	−0.090***	−0.321***	−0.339***
	(0.009)	(0.014)	(0.015)
forecasted_warning	−0.098***	−0.077***	−0.081***
	(0.006)	(0.006)	(0.006)
wrong_rain_forecast	0.021***	0.259***	0.248***
	(0.004)	(0.009)	(0.009)
wrong_dry_forecast	−0.102***	−0.161***	−0.174***
	(0.007)	(0.008)	(0.009)
forecasted_morning_rain	−0.096***		
	(0.007)		
forecasted_midday_rain	−0.053***		
	(0.008)		
forecasted_afternoon_rain	0.012		
	(0.009)		
forecasted_morning_rain_on_morning_traffic		−0.296***	
		(0.012)	
forecasted_morning_rain_on_midday_traffic		−0.159***	
		(0.010)	
forecasted_morning_rain_on_afternoon_traffic		−0.135***	
		(0.009)	
forecasted_midday_rain_on_morning_traffic		−0.015**	
		(0.007)	
forecasted_midday_rain_on_midday_traffic		−0.408***	
		(0.017)	
forecasted_midday_rain_on_afternoon_traffic		−0.101***	
		(0.010)	
forecasted_afternoon_rain_on_morning_traffic		−0.010	
		(0.009)	
forecasted_afternoon_rain_on_midday_traffic		−0.014*	
		(0.007)	
forecasted_afternoon_rain_on_afternoon_traffic		−0.363***	
		(0.015)	
forecasted_morning_rain_only_on_morning_traffic			−0.312***
			(0.013)
forecasted_morning_rain_plus_one_on_morning_traffic			−0.274***
			(0.012)
forecasted_all_day_rain_on_morning_traffic			−0.235***
			(0.011)
forecasted_midday_rain_only_on_midday_traffic			−0.305***
			(0.013)
forecasted_midday_rain_plus_one_on_midday_traffic			−0.344***
			(0.016)
forecasted_all_day_rain_on_midday_traffic			−0.431***
			(0.016)
forecasted_afternoon_rain_only_on_afternoon_traffic			−0.234***
			(0.013)
forecasted_afternoon_rain_plus_one_on_afternoon_traffic			−0.335***
			(0.013)
forecasted_all_day_rain_on_afternoon_traffic			−0.410***
			(0.015)
actual_cloudiness	−0.003*	−0.005***	−0.005***
	(0.001)	(0.001)	(0.001)
forecasted_darkness	−0.455***	−0.295***	−0.313***
	(0.019)	(0.015)	(0.016)
actual_relative_humidity	−0.008***	−0.006***	−0.006***
	(0.001)	(0.001)	(0.001)
actual_windspeed	−0.036***	−0.032***	−0.033***
	(0.002)	(0.002)	(0.002)
official_holidays	−0.780***	−0.783***	−0.783***
	(0.045)	(0.045)	(0.045)
school_holidays	−0.188***	−0.188***	−0.189***
	(0.010)	(0.010)	(0.010)
semester_break	−0.160***	−0.159***	−0.159***
	(0.014)	(0.014)	(0.014)
Observations	1,576,348	1,576,348	1,576,348
Adjusted R^2	0.783	0.785	0.784
Residual Std. Error	0.782 (df = 1576114)	0.781 (df = 1576108)	0.781 (df = 1576108)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: Dependent variable: *bicycle_counts*. Fixed effects for the bicycle counting station, hour of day, day of week, month of year, and year are included at all times. Standard errors are clustered at the level of the bicycle counting stations. It should be noted that the regression coefficients for *wrong_rain_forecast* and *wrong_dry_forecast* have both changed their sign. This is due to a change of the reference category, which is “no rain” for Regressions (4) –(7) and (11)–(14), but it is “no forecasted rain” for Regressions (15)–(17).

REFERENCES

Ahmed, F., Rose, G., and Jakob, C. (2013). Commuter cyclist travel behavior: Examination of the impact of changes in weather. *Transportation Research Record*, 2387(1):76–82.

An, R., Zahnow, R., Pojani, D., and Corcoran, J. (2019). Weather and cycling in new york: The case of citibike. *Journal of Transport Geography*, 77:97 – 112.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data*, 5.

Bergström, A. and Magnusson, R. (2003). Potential of transferring car trips to bicycle during winter. *Transportation Research Part A: Policy and Practice*, 37(8):649 – 666.

Böcker, L., Dijst, M., and Prillwitz, J. (2013). Impact of everyday weather on individual daily travel behaviours in perspective: A literature review. *Transport Reviews*, 33(1):71–91.

Cools, M. and Creemers, L. (2013). The dual role of weather forecasts on changes in activity-travel behavior. *Journal of Transport Geography*, 28:167 – 175.

Corcoran, J., Li, T., Rohde, D., Charles-Edwards, E., and Mateo-Babiano, D. (2014). Spatio-temporal patterns of a public bicycle sharing program: the effect of weather and calendar events. *Journal of Transport Geography*, 41:292 – 305.

de Chardon, C. M., Caruso, G., and Thomas, I. (2017). Bicycle sharing system 'success' determinants. *Transportation Research Part A: Policy and Practice*, 100:202 – 214.

de Hartog, J. J., Boogaard, H., Nijland, H., and Hoek, G. (2010). Do the health benefits of cycling outweigh the risks? *Environmental Health Perspectives*, 118(8):1109–1116.

Dias, G. M., Bellalta, B., and Oechsner, S. (2015). Predicting occupancy trends in barcelona's bicycle service stations using open data. In *2015 SAI Intelligent Systems Conference (IntelliSys)*, pages 439–445.

Dunkerley, D. (2008). Rain event properties in nature and in rainfall simulation experiments: a comparative review with recommendations for increasingly systematic study and reporting. *Hydrological Processes*, 22(22):4415–4435.

Ermagun, A., Lindsey, G., and Loh, T. H. (2018). Urban trails and demand response to weather variations. *Transportation Research Part D: Transport and Environment*, 63:404 – 420.

Flynn, B. S., Dana, G. S., Sears, J., and Aultman-Hall, L. (2012). Weather factor impacts on commuting to work by bicycle. *Preventive Medicine*, 54(2):122 – 124.

Gebhart, K. and Noland, R. B. (2014). The impact of weather conditions on bikeshare trips in washington, dc. *Transportation*, 41(6):1205–1225.

Handy, S. L., Boarnet, M. G., Ewing, R., and Killingsworth, R. E. (2002). How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine*, 23(2, Supplement 1):64 – 73.

Hanson, S. and Hanson, P. (1977). Evaluating the impact of weather on bicycle use. *Transportation Research Record*, 629:43–48.

Heaney, A. K., Carrión, D., Burkart, K., Lesk, C., and Jack, D. (2019). Climate change and physical activity: Estimated impacts of ambient temperatures on bikeshare usage in new york city. *Environmental Health Perspectives*, 127(3):037002–1 – 037002–10.

Heinen, E., Maat, K., and van Wee, B. (2011). Day-to-day choice to commute or not by bicycle. *Transportation Research Record*, 2230(1):9–18.

Helbich, M., Böcker, L., and Dijst, M. (2014). Geographic heterogeneity in cycling under various weather conditions: evidence from greater rotterdam. *Journal of Transport Geography*, 38:38 – 47.

Hunt, J. D. and Abraham, J. E. (2007). Influences on bicycle use. *Transportation*, 34(4):453–470.

Jorna, R. and Zoer, H. J. (2012). Safecycle: E-safety applications for safe cycling in europe. *Procedia - Social and Behavioral Sciences*, 48:589–596.

Kraemer, J. D., Zaccaro, H. N., Roffenbender, J. S., Baig, S. A., Graves, M. E., Hauler, K. J., Hussain, A. N., and Mulroy, F. E. (2015). Assessing the potential for bias in direct observation of adult commuter cycling and helmet use. *Injury Prevention*, 21(1):42–46.

Krizek, K. J. (2007). *Estimating the Economic Benefits of Bicycling and Bicycle Facilities: an Interpretive Review and Proposed Methods*, pages 219–248. Physica-Verlag HD.

Liu, C., Susilo, Y. O., and Karlström, A. (2014). Examining the impact of weather variability on non-commuters' daily activity–travel patterns in different regions of sweden. *Journal of Transport Geography*, 39:36 – 48.

Liu, C., Susilo, Y. O., and Karlström, A. (2015). The influence of weather characteristics variability on individual's travel mode choice in different seasons and regions in sweden. *Transport Policy*, 41:147 – 158.

Meng, M., Zhang, J., Wong, Y. D., and Au, P. H. (2016). Effect of weather conditions and weather forecast on cycling travel behavior in singapore. *International Journal of Sustainable Transportation*, 10(9):773–780.

Miranda-Moreno, L. F. and Nosal, T. (2011). Weather Or Not To Cycle: Temporal Trends and Impact of Weather on Cycling in an Urban Environment. *Transportation Research Record*, (2247):42–52.

Miranda-Moreno, L. F., Nosal, T., Schneider, R. J., and Proulx, F. (2013). Classification of bicycle traffic patterns in five north american cities. *Transportation Research Record*, (2339):68–79.

Motoaki, Y. and Daziano, R. A. (2015). A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand. *Transportation Research Part A: Policy and Practice*, 75:217 – 230.

Nankervis, M. (1999). The effect of weather and climate on bicycle commuting. *Transportation Research Part A: Policy and Practice*, 33(6):417–431.

Nordback, K. L. (2012). *Estimating Annual Average Daily Bicyclists and Analyzing Cyclist Safety at Urban Intersections*. PhD thesis, University of Colorado Denver.

Nosal, T. and Miranda-Moreno, L. F. (2014). The effect of weather on the use of north american bicycle facilities: A multi-city analysis using automatic counts. *Transportation Research Part A: Policy and Practice*, 66(1):213–225.

Sabir, M. (2011). *Weather and travel behaviour*. PhD thesis, Vrije Universiteit Amsterdam.

Saneinejad, S., Roorda, M. J., and Kennedy, C. (2012). Modelling the impact of weather conditions on active transportation travel behaviour. *Transportation Research Part D: Transport and Environment*, 17(2):129 – 137.

Thomas, T., Jaarsma, R., and Tutert, B. (2013). Exploring temporal fluctuations of daily cycling demand on dutch cycle paths: The influence of weather on cycling. *Transportation*, 40(1):1–22.

Tin Tin, S., Woodward, A., Robinson, E., and Ameratunga, S. (2012). Temporal, seasonal and weather effects on cycle volume: an ecological study. *Environmental Health*, 11(1):12.

Tokay, A. and Short, D. A. (1996). Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. *Journal of Applied Meteorology*, 35(3):355–371.

Winters, M., Friesen, M. C., Koehoorn, M., and Teschke, K. (2007). Utilitarian bicycling. a multilevel analysis of climate and personal influences. *American Journal of Preventive Medicine*, 32(1):52–58.

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach*. South-Western, Cengage Learning, 5 edition.

Yang, B., Lee, D. K., Heo, H. K., and Biging, G. (2019). The effects of tree characteristics on rainfall interception in urban areas. *Landscape and Ecological Engineering*, 15(3):289–296.

Zhao, J., Wang, J., Xing, Z., Luan, X., and Jiang, Y. (2018a). Weather and cycling: Mining big data to have an in-depth understanding of the association of weather variability with cycling on an off-road trail and an on-road bike lane. *Transportation Research Part A: Policy and Practice*, 111:119–135.

Zhao, P., Li, S., Li, P., Liu, J., and Long, K. (2018b). How does air pollution influence cycling behaviour? evidence from beijing. *Transportation Research Part D: Transport and Environment*, 63:826 – 838.

Westfälische Wilhelms-Universität Münster, Institute of Transport Economics, Working Paper Series

23. **“To pay or not to pay for parking at shopping malls - A rationale from the perspective of two-sided markets”**
    by Inga Molenda and Gernot Sieg, November 2017

24. **“A duopoly of transportation network companies and traditional radio-taxi dispatch service agencies”**
    by Thorsten Heilker and Gernot Sieg, November 2017

25. **“Airport Efficiency in Pakistan – A Data Envelopment Analysis with Weight Restrictions”**
    by David Ennen and Irem Batool, November 2017

26. **“Economic implications of phantom traffic jams: Evidence from traffic experiments”**
    by Kathrin Goldmann and Gernot Sieg, December 2018

27. **“Evaluating the Transport-Mode-Specific Trade Effects of Different Transport Infrastructure Types”**
    by Jan Wessel, February 2019

28. **“The impact of delays on the welfare effects of on-track competition: The case of transfer passengers with operator-tied tickets”**
    by Christina Brand and Gernot Sieg, December 2019

29. **“Ride-Hailing Services in Germany: Potential Impacts on Public Transport, Motorized Traffic, and Social Welfare”**
    by David Ennen and Thorsten Heilker, January 2020

30. **“Quantifying the phantom jam externality: The case of an Autobahn section in Germany”**
    by Kathrin Goldmann and Gernot Sieg, March 2020

31. **“TEN-T Corridors – Stairway to Heaven or Highway to Hell?”**
    by Kathrin Goldmann and Jan Wessel, June 2020

32. **“Using weather forecasts to forecast whether bikes are used”**
    by Jan Wessel, June 2020

For a complete list of Working Papers published by Westfälische Wilhelms-Universität Münster, Institute of Transport Economics, please visit the website (http://www.iv-muenster.de)