

Troxler, Pascal

**Conference Paper**

## Weather Forecasts and their Relation to Ski Demand

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Troxler, Pascal (2022) : Weather Forecasts and their Relation to Ski Demand, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/264121>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Weather Forecasts and their Relation to Ski Demand\*

Pascal Troxler<sup>†</sup>

March 1, 2022

## Abstract

In the wake of stagnating demand across Alpine ski areas, new pricing regimes and recent advances in the availability of precise local weather forecasts, the relation of weather forecasts to ski demand gains new relevance. By matching demand data of three Swiss ski areas to local forecast and weather data, I investigate how forecast errors affect ski demand above the implied weather effect. As demand patterns hinge on individual conditional weather expectations, I set up an activity choice framework that incorporates weather preferences and decision timing. I suggest that potential skiers trade off the risk of forecast errors against transaction costs. Therefore, skiers facing relatively low transaction costs, such as season-pass owners, decide on average later than those facing high transaction cost, such as one-day visitors. Using variation in weather and forecasts, I estimate the share of late deciders being larger than that of early deciders in two out of three areas. However, this estimate is likely biased towards the late deciders as stable weather situations decrease the risk of failures in optimization and allow guests to decide early without variation in forecast errors. On the contrary, the variation in forecast errors induces large effect sizes: A one standard deviation in forecast errors affects demand up to 20% while keeping the weather constant.

---

\*First draft of a forthcoming working paper.

<sup>†</sup>University of Bern, Center for Regional Economic Development, Tourism Research Unit (CRED-T)

# 1 Introduction

It is well established in the literature that weather and snow conditions have a significant impact on the ski demand (Malasevska et al., 2017a; Shih et al., 2009; Malasevska & Haugom, 2018). However, most people are likely to base their decision on weather forecasts rather than the actual weather. Weather forecast provide information about possible skiing conditions and as such shape the expectation of potential skiers. This allows a potential skier to make an early decision to engage in the activity. As actual conditions in an unstable weather environment only become truly visible at the site itself, a decision is only reconsidered after bearing the transaction costs of packing and driving to the area entrance. Therefore, the ability to judge the conditions at the site and decide close to the entrance in time and space is more bearable for those facing low transaction costs. In consequence, weather forecasts might explain demand fluctuations better than the weather itself and more so, when the forecasts deviate from the actual weather and when average transaction costs are high. Moreover, neglecting forecast errors confounds estimates of the impact of the weather itself. In the light of climate change that endangers snow-reliability in ski area operations (Koenig & Abegg, 1997; Elsasser & Bürki, 2002; Gössling et al., 2012; Gonseth, 2013; Steiger & Abegg, 2017, 2018) and is considered a key-factor in stagnating skiing demand within the Alps (Plaz & Schmid, 2015), operators react by implementing disruptive price strategies while making costly mitigation investments (Falk, 2015; Malasevska et al., 2020; Lütolf et al., 2020). Thus, accurate predictions of demand serve ski area operators in at least two aspects: It helps to plan staff and other short-term improvements on the operation that reduce daily operation cost<sup>1</sup>. And it increases pricing efficiency for those areas implementing dynamic pricing<sup>2</sup>.

Potential skiers might react to bad weather (forecasts) in three ways: First, they choose another activity, which decreases the overall level of skier days. Secondly, they can postpone their activity to a day with a better weather forecast and lastly they can switch to a different area. Exploiting daily ski demand data from three large ski areas paired with local weather and weather forecast data allows me to investigate upon the decision patterns of ski pass owners. Up to my knowledge, there is no published study<sup>3</sup> that empirically incorporates

---

<sup>1</sup>For example, what lifts to open and which slopes to groom.

<sup>2</sup>Operators increasingly base ticket prices on expected demand to extract as much as possible of the willingness to pay of skiers conditional on the season day and/or the weather (Malasevska et al., 2020; Lütolf et al., 2020). Often these pricing systems offer early bookers a large discount to essentially shift the weather risk from the area to the skier.

<sup>3</sup>I received the results of a paper by Scaglione and Doctor, 2008 that has a similar focus. The forecast data they used is outdated (based on TV forecasts) as advances in information technology make weather

the effect of local weather forecasts on skiing demand. Hence, I seek to close this gap in the literature by (i) setting up an activity choice model, (ii) identify weather and forecast preferences using machine learning techniques, (iii) recover causal estimates of optimistic and pessimistic weather forecasts on the level of skiers and (iv) adapting the framework to a dynamic context for one-week pass owners.

Surprisingly, in two out of three areas weather data is superior in predicting demand to weather forecast data. This result persists across different prediction models and validation techniques. As weather data is generated throughout the activity and not *ex ante* this result seems at first counter-intuitive. With the help of an activity choice framework, I suggest that this unexpected result has two causes. First, in the areas observed a large share of potential skiers has the possibility to decide spontaneously whether to ski or not. Moreover, potential skiers with low transaction costs (i.e. locals or overnight guests) are even incentivized to decide as late as possible due to the risk of having to ski in worse than expected conditions. The weather forecast on the eve of the day in question might deviate already by much more from the measured weather data than what the spontaneous skier observes in the morning. Consequently, weather data explains demand more accurately in this case. Secondly, it is inherently riskier to decide early in unstable large scale weather situations compared to stable situations. For instance, in a typical west-wind situation, rapid weather changes exacerbate weather forecasting relative to other situations (MeteoSchweiz, 2012). This incentivizes late decisions exactly in those instances when forecasts likely deviate from the actual weather. On the contrary, in a typical high-pressure situation clouds are suppressed and the fortunate weather persists for some days. Then, potential skiers can make risk-free early decisions but, at the same time, weather data is going to reflect these early decisions in equal measure to forecast data because of the ease to forecast precisely. In these situations the early decisions are disguised by variation in weather data only and, through that, underestimated. Aggregating these relations over unstable and stable large-scale situations, the weather data turns out to match demand in total better than the forecast data.

Yet, forecasts alter demand on top of what the weather measurements predicts. With the help of a panel data specification including individual season day fixed effects to account for the seasonality and a forecast error variable that proxies mistakes by the weather forecaster, I show empirically how these mistakes translate to a skier outcome above or below to what the conditions itself would predict. This can be explained by the supposedly smaller share of potential skiers with high transaction costs (e.g. one-day trippers). They face a trade-off

---

forecasts in recent years broader available in time (via the mobile phone) and space (tailored to municipalities or mountains).

between increasing transaction costs (e.g. late deciders might sleep in, arrive later in the area and are generally more exposed to traffic jams or finding a suitable parking spot) and decreasing weather risk over time. With large enough transaction costs, it becomes optimal to decide early based on weather forecasts. This trade-off can be partially disrupted by the availability of alternative outdoor activities that involve much lower transaction costs (e.g. biking, running, hiking etc.). In such a situation the loss invoked by a potential pessimistic forecast - that hinders the potential skier to drive to the area even though the actual weather would make this choice optimal - can be almost completely mitigated by following the alternative activity. Whereas in a situation with an optimistic forecast in which the individual bears the transaction cost by driving to the area before realizing that the weather turns out unfortunate, the loss invoked by skiing is then irreversible. In that case, the sunk cost for the individual will turn out to be too large to follow the optimal choice and stay off the mountain. The empirical analysis supports the existence of this channel by indicating this asymmetry in pessimistic and optimistic forecasts by lower effect sizes of the latter. This channel is particularly a concern for ski area operators in the wake of climate change where mild winter weather and fewer fog-days in the flatlands enhances the attractiveness of alternative outdoor activities relatively to skiing (Plaz & Schmid, 2015).

Using week-pass owners only, I analyse their weather and forecast related dynamic behaviour during their vacation. Due to their validity to ski every day, one could presume that the weather and weather forecasts have no effect on demand at all. But as week-passes have lower per-day prices, many buy them out of simple convenience and by that overtake the weather risk from the area. As the data reveals, weather and weather forecasts alter demand even when the entrance has already been paid for. Furthermore, the estimation of panel models indicate lower weather dependence to the end of week-pass owners' vacation. This could be fuelled by regret for the already paid passes (i.e. sunk cost) or due to risk polarization regarding the weather. Some skiers learn fast and raise their engagement in the activity to the end of the week or, conversely, some get tired from the required endurance and decrease their engagement. Both groups' utility of skiing changes relatively to the alternative and become less weather sensitive as consequence. Including forecast errors into the model reveals that their effect is largest mid-week, suggesting that many week-pass owners use the mid-week days to take a break when forecasts are unfortunate, or, on the other hand, want to make use of the better predicted weather after postponing from unfavourable weather at the start of the week.

Section 2 covers the activity choice model and the hypotheses that follow from it. Section 3 delves into the data, its predictive performance and the identification of weather preferences. Section 4 derives the empirical strategy to evaluate the stated hypotheses and

Section 5 covers the results before concluding in Section 6.

## 2 Theoretical Model

### 2.1 Framework

The leisure consumption choice decision is closely related to the location choice in Diamond et al., 2017, land use choice in Scott, 2014. Suppose individual  $i$  can choose activity  $\mathcal{X} = \{x = \textit{skiing}, x' = \textit{other}\}$  on any day (and thus omitting the subscript  $t$ ) and can choose at most one activity at a time. The *ex-post* utility is a function of choice, decision horizon  $h$  and measured weather  $w$ . Individuals optimally choose to ski when the utility of skiing surpasses the utility of the weather independent alternative, that is  $u(x, w, h) > u(x')$ . I assume that individuals favour good weather and are risk averse,  $\frac{\partial u(x, w, h)}{\partial w} > 0$  and  $\frac{\partial^2 u(x, w, h)}{\partial^2 w} < 0$ .

For modelling the weather and forecast errors I use a Bayesian framework originally proposed by Katz et al., 1982. The observed weather can be viewed as a random variable  $W$  with probability density function  $f_W(w)$ . Likewise, a weather forecast can be viewed as a random variable  $F$  with pdf  $f_F(f)$ . Potential skiers evaluate not only the chances of a specific weather outcome, namely the prior  $f_W(w)$ , but rather the chances of such an outcome given a certain forecast - the posterior probability distribution  $f_{W|F}(w|f)$ . I allow the prior to vary across days within a season because the weather is systematically different in early season compared to late season. But, the posterior is not allowed to vary (i.e. the accuracy of the forecast does not hinge on the day of the season, I test for this assumption in XX). In order to estimate these densities, I assume the joint distribution of  $W$  and  $F$  is bivariate normal with parameters  $\mu_W = E[W]$ ,  $\mu_F = E[F]$ ,  $\sigma_W^2 = Var[W]$ ,  $\sigma_F^2 = Var[F]$  and  $\rho = Corr(W, F)$ . Then under bivariate normality the posterior  $f_{W|F}(w|f)$  is a normal pdf with parameters

$$E[W|F = f] = \mu_W + \rho \frac{\sigma_W}{\sigma_F} (f - \mu_F) \quad (1)$$

and

$$Var[W|F = f] = \sigma_W^2 (1 - \rho^2). \quad (2)$$

The conditional expected value in (1) is linear in  $f$ . To simplify matters I assume  $\rho \frac{\sigma_W}{\sigma_F} = 1$ . Under this assumption, Equations (1) and (2) become

$$E[W|F = f] = \mu_W - \mu_F + f \quad (3)$$

and

$$Var[W|F = f] = \sigma_W^2 - \sigma_F^2. \quad (4)$$

I further assume that an update on  $F$  with a shorter time horizon  $h$  changes the variance of the posterior but not the expectation (such that  $E[F_h] = \mu_F$ ). Thus, I let the random variable  $F_h$  vary with forecast horizon  $h$ . Each update is then assumed to lower the variance by a function  $\delta(h)$  that deprecates from one to zero (i.e.  $1 \geq \delta(h) \geq 0$ ) the smaller the horizon. The posterior hinges on these continuous updates and the parameters become

$$E[W|F_h = f] = \mu_W - \mu_{F_h} + f \quad (5)$$

and

$$Var[W|F_h = f] = \delta(h)(\sigma_W^2 - \sigma_F^2). \quad (6)$$

Note as the forecast horizon diminishes to zero that the conditional variance also converges to zero. The intuition is simple: The closer the forecast horizon to the actual measurement, the more exact it is supposed to become. Once the horizon is zero the forecast is no longer a forecast but the actual measurement and as such deviates no longer. Having modelled the weather and its forecast, I turn back to the utility of a potential skier. Now, suppose that the utility contains a cost function that is decreasing in forecast horizon. Meaning that a later choice involves higher transaction costs. Then, utility is characterised as

$$\bar{u}(x, w, h) = u(x, w, h) - c(h) \quad (7)$$

with  $\frac{\partial c}{\partial h} < 0$ . The expectation at  $h$  of the ex ante utility is then evaluated against the utility of the alternative activity and the individual chooses to ski when

$$E_h[u(x, w, h) - c(h)|F_h = f] \geq E_h[u(x')|F_h = f]. \quad (8)$$

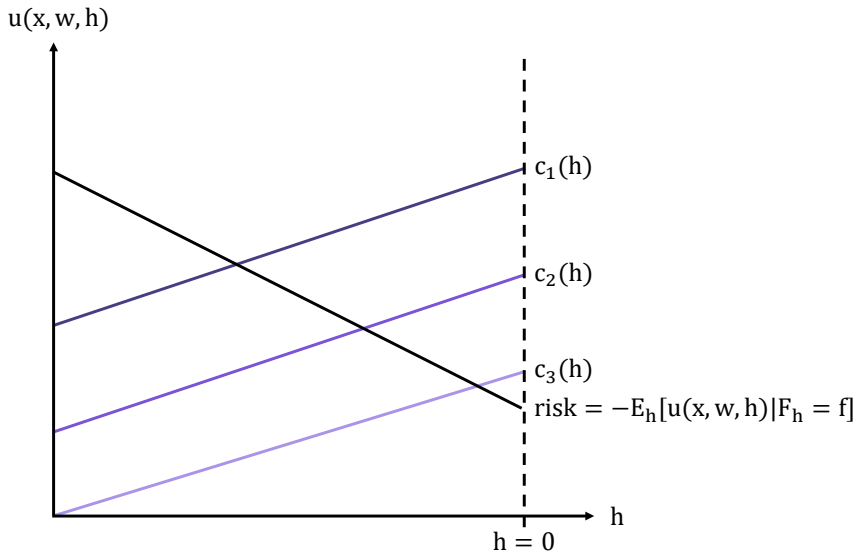
In a first step I simplify this inequality by assuming that the alternative is constant across choice horizon and weather (e.g. the individual simply rests at home and has no specific alternative in mind, then neither the weather nor the horizon at which she decides for or against skiing matters). The choice simplifies to

$$E_h[u(x, w, h)|F_h = f] - c(h) \geq K, \quad (9)$$

where  $K = E[u(x')|F_h = f]$ . Because individuals are risk averse, they face a trade-off between increasing transaction cost and decreasing risk over time. Thus  $c(h)$  is decreasing in forecast horizon  $h$  (as stated above  $\frac{\partial c(h)}{\partial h} < 0$ ), whereas the negative of the conditional expectation for a given forecast (i.e. the risk) is increasing in forecast horizon

( $\frac{\partial -E_h[u(x,w,h)|F_h=f]}{\partial h} > 0$ )<sup>4</sup>). The trade-off is depicted in Figure 1. Note that I assume one utility function across all individuals but heterogeneous cost functions - in the figure denoted as  $c_1(h)$ ,  $c_2(h)$  and  $c_3(h)$ . The situation is depicted for a given mean weather forecast that does not change over time. When the forecast horizon is still large, the variance of the forecast is still high and hinders all three individuals to decide in favour of skiing due to the risk of a bad weather outcome. The first individual to make the decision is individual 1 with the highest transaction cost. She decides at the intersection of  $c_1(h)$  and the risk curve to ski. The risk is still manageable and the transaction costs are not too high yet. It becomes clear that individuals with lower transaction cost are inclined to reduce the risk and postpone the decision to the smallest forecast horizon possible. Furthermore, note that heterogeneity in  $u(x')$  and in  $f$  are both equivalent as to shift the transaction cost curve up- or downwards.

**Figure 1: Decision trade-off over forecast horizon  $h$**



**Figure Notes:** The vertical axis indicates utility levels for a given weather forecast  $F_h = f$  and the horizontal axis the forecast horizon  $h$  with its end at  $h = 0$ . Risk ( $= -E_h[u(x, w, h)|F_h = f]$ ) is decreasing over  $h$  and depicted in black, transaction cost is heterogeneous across individuals and depicted in different shades of purple. The individual chooses to ski at the intersection of increasing individual transaction cost and decreasing risk. The lower the intercept of the transaction cost the more it is optimal to postpone the decision to a later forecast horizon.

On top of this, consider the case of risk neutral individuals: The risk curve in Figure 1

<sup>4</sup>Proof: Risk aversion implies that  $u(x, E_h[w|F_h = f], h) > E_h[u(x, w, h)|F_h = f]$ . Taking partial derivatives on both sides with respect to  $h$  and plugging in (5) yields  $\partial u(x, \mu_w - \mu_{F_h} + f, h)/\partial h > \partial E_h[u(x, w, h)|F_h = f]/\partial h$ . As the utility of skiing is independent of the decision timing itself, the left hand side is equal to 0. Therefore,  $0 > \partial E_h[u(x, w, h)|F_h = f]/\partial h$ .  $\square$



is completely flat which inclines those individuals to decide as early as possible.

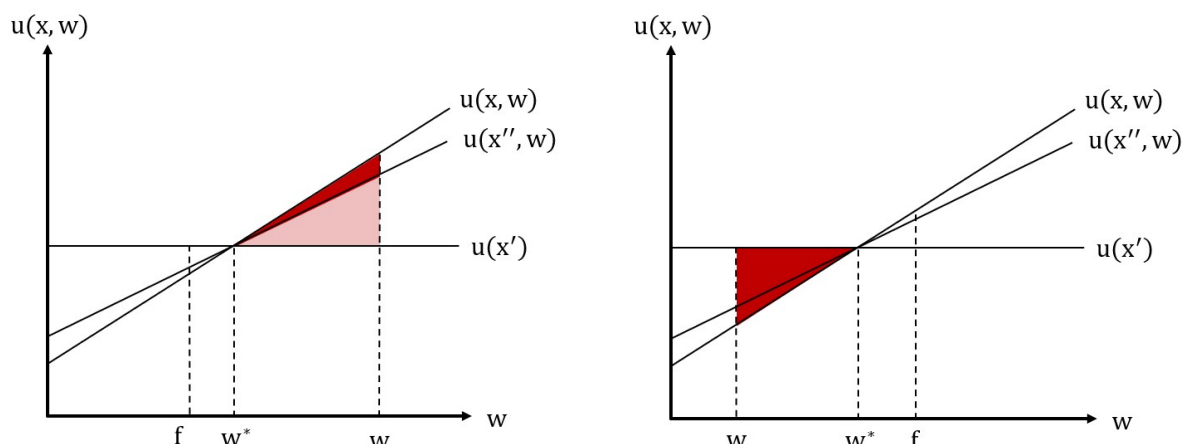
## 2.2 Three Activities

In a next step I analyse the introduction of a third activity choice. Consider a situation where next to skiing and the weather independent alternative there is a second weather dependent activity but with much lower transaction cost than skiing. One can think of cycling, running, wandering through the woods or any other low-cost outdoor activity. Now, the reason someone decides early in favour of skiing is due to its relatively high transaction cost: Having to organise all the gear, pack, get up early in the morning and drive to the mountain. If these individuals have a back-up plan that allows them to do other nice weather activities on much shorter notice, enables them to avoid losses to utility due to pessimistic forecasts. To see this, consider Figure 2. Both panels indicate risk-neutral individuals (that decide as early as possible) that fail to optimize due to wrong weather forecasts  $f$  that deviate substantially from the measured weather  $w$  and the activity indifference cut-off at  $w^*$ . In the left panel, the loss induced through the pessimistic forecast ( $f < w$ ) is almost completely offset by engaging spontaneously in the low-cost, also weather dependent outdoor activity  $x''$ . The regret (red plus light-red surface) of not being able to ski on an actually nice day is offset by the gain (light-red surface) of engaging in the alternative weather dependent activity and leaves a small loss in the size of the red surface. On the other hand, as depicted in the right panel, the loss induced through an optimistic forecast is irreversible because the individual has born the transaction cost already before learning about the true weather outcome. This person woke up early, packed all her gear and drove to the mountain before realizing that the weather is unfortunate. In that case sunk costs might be too high and the individual regretfully engages in the activity anyway. Ultimately, the possibility of mitigating losses from pessimistic forecasts and avoiding irreversible losses from optimistic forecasts increase the attractiveness of remaining home and deciding spontaneously about the activity. From this perspective, it might follow that in the aggregate the increasing availability of outdoor activities coupled with more mild temperature and less foggy winter days could partially explain the stagnation in alpine skiing demand.

## 2.3 Risk Reducing Factors

Notice that the weather risk is not static: Shorter forecast horizons and the confidence of the forecaster to predict the weather accurately can reduce these risks substantially. The actual risk reduction over forecast horizons is, surprisingly, quite small. Figure 13 in the Appendix depicts that. Although the forecast vary only slightly less over shorter horizons,

**Figure 2: Sketch of linear utility function of skiing versus weather independent alternative**



**(a) Avoiding losses in pessimistic forecasts      (b) Full losses in optimistic forecasts**

**Figure Notes:** On the horizontal axis is the weather index where  $w^*$  indicates the indifference cut-off between  $x =$  skiing,  $x' =$  not skiing and  $x'' =$  other outdoor activity. In the left panel is the weather forecast below the cut-off such that the risk neutral-type decides not to ski. The weather turns out to be nicer than the forecast ( $w > f$ ) such that skiing would be optimal. The light-red and dark-red surface combined indicate the loss in the two activity choice situation which corresponds to the avoided gain by choosing not to ski. As the potential skier can switch spontaneously to the low cost outdoor activity, almost all of this loss can be mitigated (light red surface). What remains is the red surface that indicates the actual loss between the alternative outdoor activity and skiing. In the right panel is the weather forecast above the cut-off such that the risk neutral-type decides to ski. Unfortunately, the weather turns out to be much worse at  $w < f$  and imposes an irreversible loss on the individual as transaction costs are borne already. Note that a risk-averse individual with high transaction costs faces a similar situation bar the linearity of the utility functions.

most individuals might still perceive shorter horizons to be substantially more accurate. This might be one reason why predicting demand works better with smaller forecast horizons (see Figure 5 in Section 3.4).

The more prevalent risk-reducing factor is the confidence of a weather forecast. To see this, consider two common large-scale weather conditions: A high pressure zone and a typical west-wind situation. In a high pressure zone the weather is sunny, stable and often persists for a couple of days. Consequently, forecasts during such conditions are typically more accurate compared to other large-scale weather conditions like a west-wind situation. In such conditions, the weather can completely reverse within hours in the northern Alps (MeteoSchweiz, 2015). The forecaster, knowing this, is less confident in her own forecast and communicates this in one way or the other.<sup>5</sup> In consequence, large-scale weather conditions might lead to different confidence-levels that a potential skier attributes to the forecast and

<sup>5</sup>In a TV-show the host uses terms like unstable or changing weather or in a local online-forecast it might be visible by varying forecasts for the same day within hours. Later on, I use the difference between two forecasts for the same day as a proxy of confidence in a given forecast.

this, in turn, relates directly to the risk behaviour of that individual. Small confidence bands reduce the *ex ante* risk to the individual and incentivize earlier decisions compared to large confidence bands.

The ability of the forecaster in stable versus unstable large-scale weather situations is depicted in Figure 3. Stable weather situations are high pressure or “Bise”<sup>6</sup> situations. Unstable weather situations are defined as west-wind situations, south- and north-“Foehn”<sup>7</sup> and in-between situations as well as unclear situations. Comparing the root-mean-squared error for both situations shows that in the key-variables of precipitation and sunshine weather risks are reduced by a lot. The precipitation RMSE in area 1 is roughly 25 times higher in unstable large-scale situations. For the sunshine variable forecasts are more than 2.5 times less accurate in unstable situations. However, the minimum temperature is actually less predictable in stable situations which diminishes the risk reduction in the overall weather index. In area 2 the risk enhancing temperature predictability translates to an equal weather risk for stable and unstable situations.

The risk reduction in deciding early through stable weather situations is exactly what we observe in the aggregate in Section 3.4. In unstable situations the risk of deciding early is high. Which means that individuals are incentivized to decide spontaneously upon the observed weather instead of the forecast. Hence, predicting aggregate demand in such situations should be more accurate using weather instead of forecast variables. On the contrary, the risk of deciding early is low in stable situations. Thus, individuals are incentivized to decide early to reduce transaction costs as much as possible. In such situations, it would be naturally to assume that forecasts predict demand better than the actual weather. However, as forecasting is much easier in stable weather situations forecasts deviate much less from the weather. And, thereby, demand predictions using forecast or weather data turn out to be equally precise even though people should, at least in theory, decide early in stable situations. It is not surprising then, that predicting demand over all weather situations leads to superior results using weather variables instead of forecast variables. Exactly what I observed in Section 3.4.

## 2.4 Dynamic Case

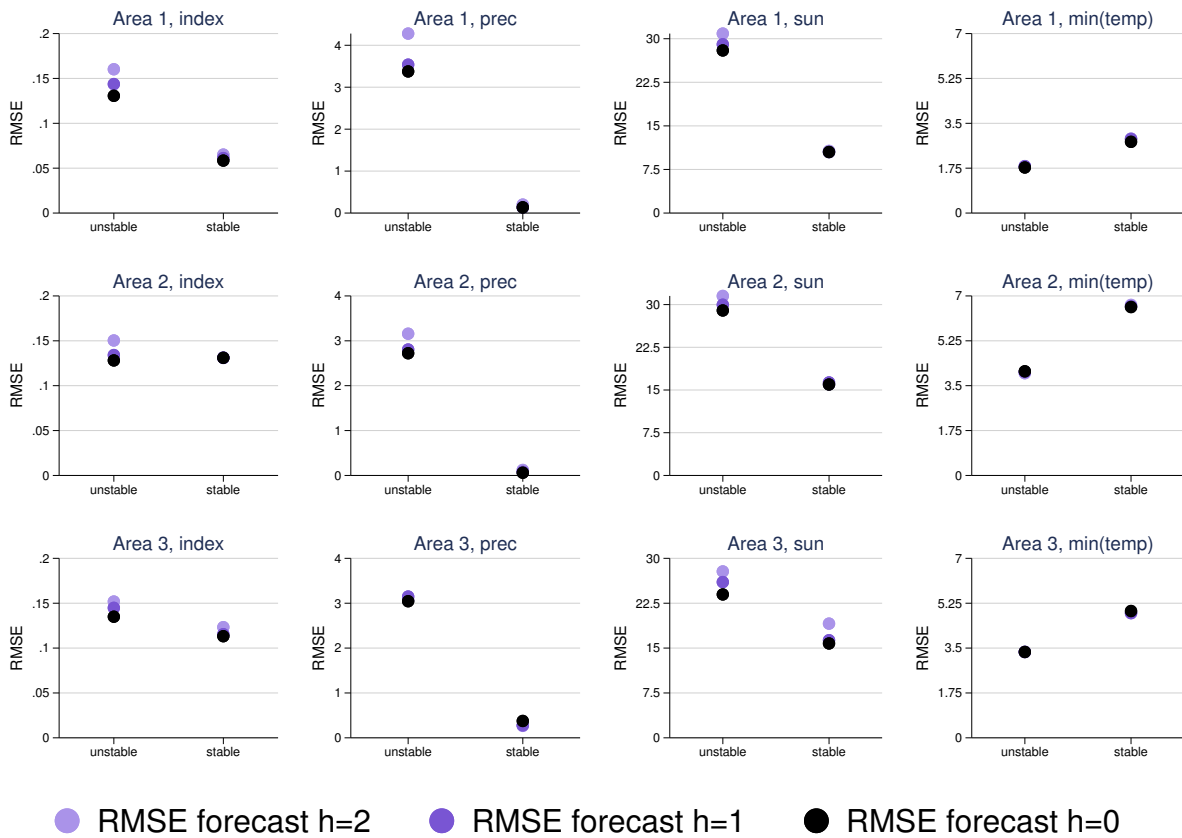
I model the dynamic activity choice in close relation to the static case in Section 2.1. Owing to the aggregate data, following a one-day or season pass owner on a dynamic path is

---

<sup>6</sup>east or nord-east winds that are channelled through the flat lands and are often accompanied by a fog cover in the lower altitudes

<sup>7</sup>strong warm winds in either northern or southern alpine valleys that are accompanied by a cloud cover on either the northern or southern Alps. In those situations the cloud cover in inner-alpine regions is hardly predictable

**Figure 3: RMSE of forecast horizons split by large-scale weather situation, areas and weather variables**



**Figure Notes:** On the horizontal axes is an indicator of stable weather situations depicted and the vertical axes indicate the root-mean-squared errors (RMSE) of different weather variables. Stable weather situations are high pressure systems and “Bise” situations (with an inversion at the fog cover and east to north-east wind system (MeteoSchweiz, 2015)). The RMSE are depicted for all three forecast horizons  $h = 1, 2$  or  $3$ . In each row is an area displayed starting from weather indices in the first column and in columns 2 to 4 the three key-weather variables identified in Section 3.5.

infeasible. Hence, the dynamic choice is best analysed by the behaviour of one-week pass owners. Albeit being less weather dependent, their sensitivity to the weather (see Figure 7) implies that some one-week pass owners do not fully use up all valid area entrances. This is clearly visible in the two gray highlighted Tuesdays in Figure 4. Noticeable is also the seasonality of these individuals that use their passes much less on the arriving or departing weekends.

For simplicity, suppose individual  $i$  has very low transaction cost (which is plausible for weekly overnight guests) and chooses activity  $\mathcal{X} = \{x = \textit{skiing}, x' = \textit{other}\}$  by consulting the actual weather  $w$  (and not forecasts) and the quantity of previous consumption  $q_x$  and  $q_{x'}$ . Next to favouring good weather  $\frac{\partial u(x,w,q_x)}{\partial w} > 0$ , individual utilities are increasing in  $q_x$  and

$q_{x'}$ . I make no prior assumption about a potential skiers marginal utility to the quantity of consumption. As the individuals bought their ski passes already, the individuals evaluate each day in their pass validity spell whether  $u(x, w, q_x) > u(x', q_{x'}) - r(q_{x'})$ . The last term is a regret function that is increasing in  $q_{x'}$  (i.e.  $\frac{\partial r(q_{x'})}{\partial q_{x'}} > 0$ ). It models the regret of not using the pass which increases by the number of unused days.

## 2.5 Hypotheses

From the Sections 2.1 to 2.3 I derive three hypotheses that are empirically investigated:

- (i) Wrong weather forecasts alter demand above or below the counterfactual situation with an accurate forecast.
- (ii) Groups that are expected to have relatively higher transaction costs have a higher share of early deciders (e.g. one-day trippers and/or families).
- (iii) Optimistic forecasts have lower effect sizes than pessimistic forecasts.

Hypothesis (i) relates directly to the decision timing of individuals. If the hypothesis is confirmed, then it follows that some individuals decide early due to increasing transaction costs or risk-neutrality and thus fail to optimize in wrong forecast situations. In hypothesis (ii), I check whether individuals in areas with a large share of one-day trippers and/or families - with arguably higher individual transaction costs - fail more often to optimize due their incentive to decide earlier. Hypothesis (iii) follows from the static case with three activities. As the individual's loss is irreversible in uncertain weather situations when forecasts predict good weather (and as consequence might be optimistic) whereas individuals can mitigate losses from pessimistic forecasts by switching to a third weather dependent activity, the risk of engaging in skiing in potential optimistic situations is higher than the risk of not skiing in a potential pessimistic situation. Therefore, I conjecture that individuals incorporate this and decide later in potential optimistic situations than in potential pessimistic situations. From this follows directly hypothesis (iii). However, one has to be cautious that a confirmation in hypothesis (iii) does not suffice to prove the above behaviour on an individual level. It serves simply as descriptive evidence that such a behaviour might be at the bottom of the observed relationship.

In a dynamic setting with two choices, I test the following hypotheses empirically:

- (I) There is satiation in individual utilities when skiing is consumed repeatedly.
- (II) Skiers become less weather dependent to the end of their stay.

(III) Forecast errors affect skiers less at the end of their stay.

Hypothesis (I) simply states that the individual’s marginal utility with respect to repeated skiing is decreasing. The idea is that guests have a preference to experience other activities than skiing alone and, thus, the marginal utility of skiing is decreasing in the quantity of its consumption (i.e.  $\frac{\partial^2 u(x,w,q_x)}{\partial^2 q_x} < 0$ ). Hypothesis (II) is directly related to the regret function  $r(q_x)$ . At the end of the week the regret for those leaving out some days becomes large enough that they ski under worse weather conditions than one would expect. Lastly, Hypothesis (III) is similar to (II) and simply tries to confirm or reject patterns from (II).

## 3 Data

### 3.1 Demand Data

Demand data consists of either bookings or first-entries<sup>8</sup> of three ski areas during winter season (end of November until the end of April). Data is available for different age groups and numerous pass validity durations. It is provided by mountain railway operators located in the western Alps of Switzerland. In total the data consists of 3302 days split into 910, 1137 and 1255 days from area 1, 2 and 3, respectively, covering all seasons between 2010 and 2020. Unfortunately, demand in area 1 is restricted to transactions, not the actual consumption of skiing. To be more certain that a transaction leads to consumption, I only use bookings of one-day passes that are valid on the same day as the transaction.<sup>9</sup> The validity and age groups provide insights about the heterogeneity in behaviour to weather and weather forecasts. To make use of these heterogeneities, data has to be aligned in a comparable fashion between the areas. Therefore, data is aggregated to three age-groups: Adults, juveniles and children, and five pass validity types: One-day passes, weekend passes (2-4 days), one-week passes (5-7 days), two-week passes (8-14 days) and season passes (more than 15 days).

### 3.2 Weather Data

Weather data is available from the Swiss government meteorological service MeteoSwiss. The data can be drawn from any of the approximately 2700 weather stations in Switzerland. The prior weather variables that are of main interest, according to the following literature, are:

---

<sup>8</sup>First-entries are daily counts of guests entering a ski area.

<sup>9</sup>It is likely that buying a seven-day pass does not necessarily resolve in seven days of actual consumption. On the contrary, it is very unlikely that buying a one-day pass leads to no consumption.

- Sunshine: This variable is the percentage share of sunshine-hours on a given day of the maximum possible sunshine-hours. Thus, it allows to compare sunny days independent of their timing within the year. An increase in sunshine duration is associated to counteract cold temperatures, having clear vision of the slope and the surroundings and facilitate consumption of lunch and other beverages on the mountain. Therefore, I expect a positive relation to skiing demand. (Gonseth, 2013; Malasevska et al., 2020; Lütolf et al., 2020; Scaglione & Doctor, 2008; Rutty & Andrey, 2014; Haugom & Malasevska, 2019).
- Precipitation: The overall sum of precipitation during a day in millimetre. How it translates into snow is mostly dependent on temperature and wind which is not relevant for this purpose. See the fresh snow variable for a further discussion. During snowfall or rain, the light is often flat which exacerbates skiing and other activities on the mountain. Often it is accompanied by cold temperatures and stormy winds. That is why I expect a negative relation to skiing demand. (Falk, 2015; Malasevska et al., 2020; Scaglione & Doctor, 2008; Rutty & Andrey, 2014; Haugom & Malasevska, 2019).
- Wind Chill Temperature: A combination of wind and temperature gives the Wind Chill Temperature (WCT). This variable indicates the perceived temperature and as such adds wind to the equation which decreases the WCT as it gets stronger. Temperatures that are too warm wets the snow that is perceived negatively by many skiers but, on the contrary, too cold temperatures might become unbearable for a large share of skiers. (Osczevski & Bluestein, 2005; Malasevska et al., 2017b)<sup>10</sup>.
- Fresh Snow: Snowfall is already measured in the precipitation variable. Thus, the precipitation of snow is shifted one day into the future to represent the fresh snow fallen within the last 24 hours. Thus, the precipitation variable depicts immediate snowfall or rain. Fresh

---

<sup>10</sup>A larger body of literature uses direct temperature measures, e.g. Shih et al., 2009; Gonseth, 2013; Falk, 2015; Scaglione and Doctor, 2008; Holmgren and McCracken, 2014

snow can have a different effect on demand than actual precipitation. The former is sought by freeride-enthusiasts whereas the latter is generally considered as being bad for skiing - clouds dim the light and the falling snow or rain enhances the bad sight further. (Shih et al., 2009; Gonseth, 2013; Falk, 2015; Holmgren & McCracken, 2014)

Before using the above weather variables in the analysis, I evaluate empirically whether these are the core variables that explain skier demand in my data. More specifically, I draw upon a broad set of additional weather variables and run statistical learning procedures, like random forests (RF), to predict skier demand and evaluate the variables that are key to the prediction outcome. This procedure is described in Section 3.5. The weather variables additionally included are: Daytime minimum, maximum and mean Temperature [°C], daily average wind speed [m/s], daily maximum wind speed (from hourly averages) [m/s], daily maximum in gusts [m/s], wind direction [°], relative humidity (daily minimum, mean and maximum) [%] and daily average air pressure [hPa]. As the weather data is not measured exactly at the same coordinates as the weather forecast model outputs, the data requires an interpolation from several weather stations. I apply inverse distance weighting on weather stations within 30km of the corresponding highest and lowest lift stations in the areas. These are exactly the same coordinates whereof weather forecast outputs from MeteoSwiss are available. For detailed information on the spatial join of the weather stations to the lift coordinates read Appendix 8.1.

### 3.3 Weather Forecast Data

The weather forecast data are midnight model output of COSMO-7 (MeteoSchweiz, 2012) and are provided by MeteoSwiss. Forecast data are available for three time-horizons, 2 days ( $f_t^{t-2}$ ), 1 day ( $f_t^{t-1}$ ) and 0 day ( $f_t^{t-0}$ ) in advance. Additionally, the data cover daily relative sunshine duration [%], daytime temperature (minimum, maximum, mean and wind chill) [°C], daily wind speed (mean, hourly minimum and maximum) [m/s], daytime precipitation [mm] and daily fresh snow [cm]. The variables are geographically referenced to the point of the lowest lying lift within all three areas such that they match the weather data in space (see Appendix 8.1 for details). One limitation of these data are that they are computed outputs that must not necessarily represent actual published forecasts perfectly. But, as the data are available on a very local scale, these are exactly the inputs forecaster use for publishing local forecasts in their mobile phone application or online. The variation in these data is likely representing actual variation in forecasts above the publication of MeteoSwiss.



Several other weather service providers draw upon the same computer model outputs to publish their forecasts. A second limitation is that the available snowpack data is of poor quality. Especially in the early season the natural height of snow might be a decisive factor of the intention to ski. To account for this without reliable snow data, I dropped observations before mid-December. At days prior to this cut-off, demand data varies much more across seasons than in later days and might potentially confound the results. (actually included in the analysis)

### 3.4 Demand Prediction

The following analysis has the purpose to reveal what share of the variation in demand can be predicted using weather and forecast variables. For this purpose we estimate the following specification:

$$y_{ds} = w'_{ds}\beta + \alpha_d + e_{ds}\gamma + \varepsilon_{ds}, \quad (10)$$

where  $y$  is demand at day  $d$  in season  $s$ ,  $w$  is a vector of the respective weather or forecast variables,  $\alpha_d$  is the season day fixed effect,  $e$  is a dummy indicating Easter holidays and  $\varepsilon$  is the error term. The vector of weather (forecast) variables is:  $w' = \{freshsnow, sunshine, precipitation, precipitation, wct\}$  where  $wct$  stands for Wind Chill Temperature. This measure represents the temperature that a person actually feels after incorporating the wind<sup>11</sup>. Exploiting the same weather and forecast variables allows a comparison of the within sample predictive performance of both variable types directly. The season day fixed effect is included to capture the entire seasonality of demand, to control for the confounding effect on the weather<sup>12</sup> and to remove any serial correlation between successive days. Controlling for this, it is reasonable to assume that the weather and forecasts are exogenous to demand. In practice, the fixed effect is computed relative to the same Saturday of each season opening. This means that the opening day and all following days are synchronized across seasons, independent of their actual date<sup>13</sup>. Note that this specification hinges on the assumption that weather-related effects of season days on demand remain constant over time.<sup>14</sup> The weather and forecast variables included are derived from the literature (see Section 3.2) and availability of forecast variables. I estimate (10) in levels by two estimators that yield equivalent weather coefficients using one-day passes only. Surprisingly, applying the least-squares dummy variable

---

<sup>11</sup>The Wind Chill Temperature (WCT) is defined as  $wct[C] = 13.12 + 0.6215T - 11.37V^{0.16} + 0.3965TV^{0.16}$  with temperature  $T$  in  $C$  and wind speed  $V$  in  $km/h$ .

<sup>12</sup>For instance temperatures are systematically higher in March than in January.

<sup>13</sup>The season openings are often in mid-December, but rarely ever at the exact same date.

<sup>14</sup>E.g. Artificial snowmaking facilities or new lifts that reach a higher elevation could lower dependence on natural snow. In Appendix 8.2 I include season fixed effects  $\varsigma_s$  to check upon its confounding effect of different weather variables.

(LSDV) and the within (WT) estimator to the same model does not necessarily imply equal prediction accuracy. The reason for this lies in the bias variance trade-off in predictions (Hastie et al., 2008). The WT estimator demeans all regressors by season days and results in exactly the same  $\hat{\beta}$  as in the LSDV estimator.<sup>15</sup> In contrast to the LSDV estimator, the demeaning purges the between season day variation before estimating (10). Predicting demand using the WT estimator requires therefore only the demeaned weather variables as regressors compared to the additional requirement of 155 dummy variables by the LSDV. The higher degrees of freedom of the WT estimator produce mechanically a worse fit. This results in a much lower variance at the expense of some bias in the mean squared error (MSE) of the prediction (see e.g. Hastie et al. (2008)). A more detailed explanation of this is in Appendix 8.3.

Table 1 indicates two measures of prediction performance, the  $R^2$  recovered from the linear regression in (10) (with level demand as outcome) and the measure  $1 - \frac{RMSE}{SD}$  that proxies out-of-sample prediction performance and is recovered from 10-fold cross-validation of the same models. In the last two columns of Table 1 and in Figure 5 the consequence of the variance-bias trade-off is clearly visible. The LSDV estimator has a much higher in-sample fit, visible by the  $R^2$  around 0.7 in area 1 and 3. It fits the data very tight due to its 164 regressors, whereas the WT has quite a low in-sample fit ( $R^2 \approx 0.18$  in area 1 and 3) with its 9 regressors. On the contrary, the WT estimator leads to much preciser out-of-sample predictions, seen by the measure  $1 - \frac{RMSE}{SD}$ . The WT estimator predicts, as seen in Table 1, up to 46% of the variation in demand correctly while the LSDV estimator achieves only up to 37%. In the first two columns of Table 1 it is additionally visible that the season days explain a larger part of the variation of ski demand than the weather itself. Regressing demand on season days or the weather alone and comparing the training set fit of the two reveals a much larger  $R^2$  for the season day fixed effects alone. Again, this is unsurprising because of the larger amount of regressors. Yet, the prediction accuracy using only season day effects remains above the weather effects alone and is roughly twice as accurate. Therefore, the seasonality plays a larger role in demand fluctuations as the weather. The weather alone is estimated to explain demand fluctuations up to 25% which is close to what the previous literature on the topic says (see e.g. Plaz and Schmid, 2015 that estimate around 15%).

**Table Notes:** The table depicts the in-sample fit as  $R^2$  and a proxy for the out-of-sample fit as  $1 - \frac{RMSE}{SD}$  (one minus the root-mean-squared error divided by the standard deviation of demand) recovered from 10-fold cross validation across the four models indicated by the columns and across the three areas indicated by the rows. The sample consists of one-day passes only. Column 1 represents a model using season day fixed effects and the Easter dummy only, column 2 using the 7 weather variables in (10) only and columns three and four represent the whole model but different estimators. The third column depicts the model fit of the

---

<sup>15</sup>By the FWL-theorem, see e.g. Lovell (2008).

**Table 1: Training and test model fit**

		s. day FE only	weather only	LSDV full	WT full
area 1	in-sample fit ( $R^2$ )	0.521	0.225	0.690	0.179
	out-of-sample fit ( $1 - \frac{RMSE}{SD}$ )	0.189	0.096	0.322	0.437
area 2	in-sample fit ( $R^2$ )	0.475	0.195	0.579	0.115
	out-of-sample fit ( $1 - \frac{RMSE}{SD}$ )	0.182	0.099	0.253	0.345
area 3	in-sample fit ( $R^2$ )	0.527	0.252	0.711	0.188
	out-of-sample fit ( $1 - \frac{RMSE}{SD}$ )	0.213	0.137	0.367	0.459

least-squares-dummy variable estimator and the fourth column the model fit of the within estimator.

Figure 5 shows - on top of the better prediction accuracy of the WT estimator compared to the LSDV estimator - that the shorter the forecast horizon, the better the demand predictions in area 1 and area 3. Only in area 2 the 0-day and 1-day forecast exceed the weather data in prediction accuracy across both estimators. An explanation for this is that the average individual among the weather sensitive one-day pass owners decides earlier upon the activity in area 2 than in areas 1 and 3. Note that I estimate overall fewer weather sensitive types in area 2 as the prediction accuracy of weather and forecasts are generally lower compared to the other areas. This explanation is further discussed in the static activity choice model and its empirical implementation in Section ??.

### 3.5 Identification of Key Variables

In this section random forests (RF) are applied to improve upon variable selection. The idea is to find the key variables explaining skiing demand to create a weather and forecast index that proxy actual and expected skiing conditions in a single variable. Because there is no built in structure in RF to account for heterogeneity, the prediction accuracies result in the range of the LSDV. But the performance of the RF is not of interest here.

First, I grow a RF on (10) including the full set of available weather variables  $w$  listed in Section 3.2 in addition to all season day dummy variables. Then, the out-of-bag (OOB) root-mean-squared error (RMSE) rate for the RF is computed. This validation technique uses part of the data as training set and the remaining data as test set to validate the prediction accuracy of the RF. The RF grows random trees and uses for each tree around two-thirds of the bootstrapped observations. On top of that, it tests the accuracy with the remaining third of observations that were not used in growing the tree. If the number of trees grown is sufficiently large, the OOB can be shown to be equivalent to the leave-one-out

cross validation (LOOCV) (Hastie et al., 2008).

The OOB is also used to estimate variable importance. By randomly permute a variable in the OOB sample, comparing the MSE to the one obtained by the original variable and average this over all trees, the decrease in the prediction accuracy of each variable is measured (Hastie et al., 2008). Another option would be to use the Lasso estimator and evaluate which parameters of weather variables are shrunk to zero at latest (while increasing the shrinking parameter  $\lambda$ ). Because the dataset includes many potential predictors that are highly collinear (as e.g. maximum and minimum temperature) OLS and likewise Lasso might struggle in separating the effects of the collinear variables. The variable importance measure computed by RF does not suffer from this limitation. Hence, I refrain from other methods to compute variable importance.

Even though the prediction performance of this RF is very inaccurate (it predicts only around 0.2 of the variation in demand accurately), it delivers valuable estimates of variable importance. Figure 6 indicates undoubtedly the three variables that contribute the most to the reduction in the MSE: The relative sunshine, the minimum temperature through the day and precipitation. After these, there is no clear ordering of variables apparent when comparing the areas. Some measure of wind and humidity have predictive power too but at a much lower magnitude than the first three.

In line with the results from the variable importance measure, I construct two indices that proxy skiing conditions by the use of weather and forecast variables of relative sunshine, minimum temperature and precipitation. First I define partial indices

$$\begin{aligned} sunInd &= sun \\ precInd &= \begin{cases} 100 - (prec * 10), & \text{if } prec \leq 10 \\ 0, & \text{otherwise} \end{cases} \\ tempInd &= \begin{cases} 100 - (|-10 - temp|) \cdot 5, & \text{if } |temp| \leq 20 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

that range for each variable from 0 (worst condition) to 100 (best condition). For example a relative sunshine forecast of 0% is the worst possible weather forecast concerning sunshine duration. Whereas a sunshine forecast of 100% provides the best conditions. The worst conditions for precipitation and temperature forecasts are at least 10mm precipitations and a deviation from -10°C by at least 20°C (i.e. -30°C or +10°C). In between those values, the conditions increase linearly to the best case of 100. In a next step I weight each partial index

in accordance with the variable importance measures<sup>16</sup> to construct the overall weather and forecast indices

$$w = 1/3 \cdot sunInd + 1/3 \cdot precInd + 1/3 \cdot tempInd$$

that can equally take values from 0 to 100. It is important to note that all results presented in the following sections are robust to different weightings and boundaries of these three variables.

## 4 Empirical Strategy

### 4.1 Static Case

The first empirical model to test hypotheses (i) and (ii) is derived from the individual *ex ante* utilities in Equation (??). I estimate

$$\log(y_{ds}) = w_{ds}\beta + \Delta_{ds}\delta + \alpha_d + e_{ds}\gamma + \varepsilon_{ds}, \quad (11)$$

where  $w_{ds}$  is the weather index,  $\Delta_{ds}^{d-h} = f_{ds}^{d-h} - w_{ds}$  is the forecast error,  $\alpha_d$  is the season day fixed effect that is common across seasons but fixed across season days,  $e$  is a dummy indicating Easter holidays and  $\varepsilon$  is the error term.

Let us derive the identification strategy of (11) from the average of all individual *ex ante* utilities in (??). On a given day  $d^*$   $N^A$  engage in skiing independently from the weather,  $N^L$  individuals decide late using  $w_t$  in *ex ante* utility  $\bar{u}(w_t)$  and  $N^E$  decide early using  $f_t^{t-h}$  at a given  $h$  (and thus neglecting the superscript). Then the average of all individual *ex ante* utilities is

$$\bar{\bar{u}}_{d^*} = \frac{1}{N} \left[ \sum_{i=1}^{N^A} (\eta_{id^*} + \lambda_{id^*}) + \sum_{i=1}^{N^L} (\gamma_i w_{d^*} + \eta_{id^*} + \lambda_{id^*}) + \sum_{i=1}^{N^E} (\gamma_i f_{d^*} + \eta_{id^*} + \lambda_{id^*}) \right]. \quad (12)$$

Substituting all elements for their sample averages and the forecast  $f_{d^*}$  for the difference of

---

<sup>16</sup>Not exactly by their share of MSE increase but rather simply by one third

the weather and the forecasts deviation  $f_{d^*} = w_{d^*} + \Delta_{d^*}$  yields

$$\bar{u}_{d^*} = \frac{N^{\mathbf{A}}}{N} (\bar{\eta}_{d^*} + \bar{\lambda}_{d^*}) + \frac{N^{\mathbf{L}}}{N} (\bar{\gamma}w_{d^*} + \bar{\eta}_{d^*} + \bar{\lambda}_{d^*}) + \frac{N^{\mathbf{E}}}{N} (\bar{\gamma}w_{d^*} + \bar{\eta}_{d^*} + \bar{\lambda}_{d^*}) + \frac{N^{\mathbf{E}}}{N} \bar{\gamma} \Delta_{d^*} \quad (13)$$

$$= \underbrace{\bar{\eta}_{d^*} + \bar{\lambda}_{d^*}}_{=\alpha_{d^*}} + \underbrace{\frac{N^{\mathbf{L}} + N^{\mathbf{E}}}{N} \bar{\gamma}}_{=\beta} w_{d^*} + \underbrace{\frac{N^{\mathbf{E}}}{N} \bar{\gamma}}_{=\delta} \Delta_{d^*}. \quad (14)$$

By comparing (11) and (14) the underlying assumptions in (11) become straight-forward. The outcome is a proxy of the average *ex ante* utility across individuals: The more skiers visit an area the more individuals choose to ski due to their expected gain from utility. The average individual taste parameter  $\bar{\lambda}_{d^*}$  and the average individual fixed effect  $\bar{\eta}_{d^*}$  are subsumed under the season day fixed effect  $\alpha_{d^*}$ . The  $\beta$  in the aggregate specification captures the average effect of the weather weighted down by the weather and forecast dependent groups  $\mathbf{E}$  and  $\mathbf{L}$ . Whereas the coefficient on the forecast error  $\delta$  captures the average effect of a forecast error weighted down by the fraction of the  $\mathbf{E}$ -type group. The smaller these groups' sizes are the lower are the effect sizes of the weather and forecast errors in the aggregate. This is exactly why in hypothesis (i) the effect sizes should be larger for areas more abundant in  $\mathbf{L}$ -types and in hypothesis (ii) the effect sizes on the forecast error should be higher for areas more abundant in  $\mathbf{E}$ -types. Suppose all skiers are of type  $\mathbf{E}$ , i.e.  $N^{\mathbf{E}} = N$ : A forecast error reverses the impact of the weather completely. For instance, when the actual weather is perfect, implying  $w_d = 100$ , but the forecaster predicted the worst possible weather with a  $\Delta_d = -100$ , then only those engage in skiing that prefer skiing over the alternative even in the worst conditions. It leads to the same outcome as if the weather was actually at its worst possible value  $w_d = 0$  but accurately predicted. On the other hand, suppose all skiers are of type  $\mathbf{L}$ : The last term of (14) drops out and the  $\delta$ -coefficient in the estimation of (11) is indistinguishable from zero. Hence, in the case of everyone deciding late the weather forecast becomes redundant in explaining skier outcome. A convenient property of estimating (11) is that the estimates  $\hat{\beta}$  and  $\hat{\delta}$  allow to recover the fraction of  $\mathbf{E}$ -types from the weather sensitive types. Using (14) it becomes clear that  $\frac{N^{\mathbf{E}}}{N^{\mathbf{E}} + N^{\mathbf{L}}} = \hat{\delta} / \hat{\beta}$ . Additionally, note that differences in  $\hat{\beta}$  across areas, ticket- or age-groups can stem from either the fraction of weather sensitive types (i.e.  $\frac{N^{\mathbf{L}} + N^{\mathbf{E}}}{N}$ ) or from the individuals' average elasticity with respect to the weather (i.e.  $\bar{\gamma}$ ). Unfortunately, these two effects cannot be disentangled from the data.

## 4.2 Dynamic Case

The answer to these hypotheses is best analysed by looking at one-week pass owners. Due to the fact that most people go on their vacation weekend to weekend, also most bookings

of one-week passes start at some point during the weekend.<sup>17</sup>

To evaluate the dynamic activity choice, the decision timing (late versus early decision) is, as explained above, neglected first. I simply differentiate between weather independent deciders ( $N^{\mathbf{A}}$ ) and weather sensitive deciders ( $N^{\mathbf{W}} = N^{\mathbf{E}} + N^{\mathbf{L}}$ ). Using (??), neglecting (first) that  $\gamma$  might change over consumption and taking averages over week-day  $d^*$  yields

$$\bar{u}_{d^*} = \frac{N^{\mathbf{A}}}{N} (\bar{\phi}_{d^*q_x}) + \frac{N^{\mathbf{W}}}{N} (\bar{\gamma}w_{d^*} + \bar{\phi}_{d^*q_x}) \quad (15)$$

$$= \underbrace{\bar{\phi}_{d^*q_x}}_{=\delta'_d} + \underbrace{\frac{N^{\mathbf{W}}}{N}\bar{\gamma}}_{=\beta}w_{d^*}. \quad (16)$$

The underbraces in (16) indicate the identification using aggregate data. The model to be estimated is

$$\log(y_{\tilde{d}ks}) = \beta w_{\tilde{d}ks} + \delta' \tilde{d} + \alpha_k + e_{\tilde{d}ks} + \varepsilon_{\tilde{d}ks}, \quad (17)$$

where  $\tilde{d}$  indices the week-day 0 – 6 for Sunday to Saturday, respectively,  $k$  the weeks that are lined up parallel to the season days  $d$  and  $s$  indicate, as above, seasons.  $\alpha_k$  is the week fixed effect that serves the same purpose as the season days in the baseline model<sup>18</sup>,  $\beta$  is the coefficient on the weather elasticity,  $\delta_d$  is the vector of coefficients of the week-day effects,  $e_{\tilde{d}ks}$  is again the Easter-holiday dummy. The parameters of interest are the week-day effects that translate one-by-one to the parameter  $\phi_{dq_x}$  as shown in (16).

---

<sup>17</sup>Separating booking and demand data by starting weekdays reveals that validity spells of 7-day passes start mostly on Saturdays, of 6-day passes on Sundays and of 5-day passes on Mondays. The See Figure 15 and 14 in Appendix 8.7

<sup>18</sup>The week fixed effects account for the seasonality of the different weeks and correct for the fact that the weather in late season might systematically differ from the early season.

## 5 Results

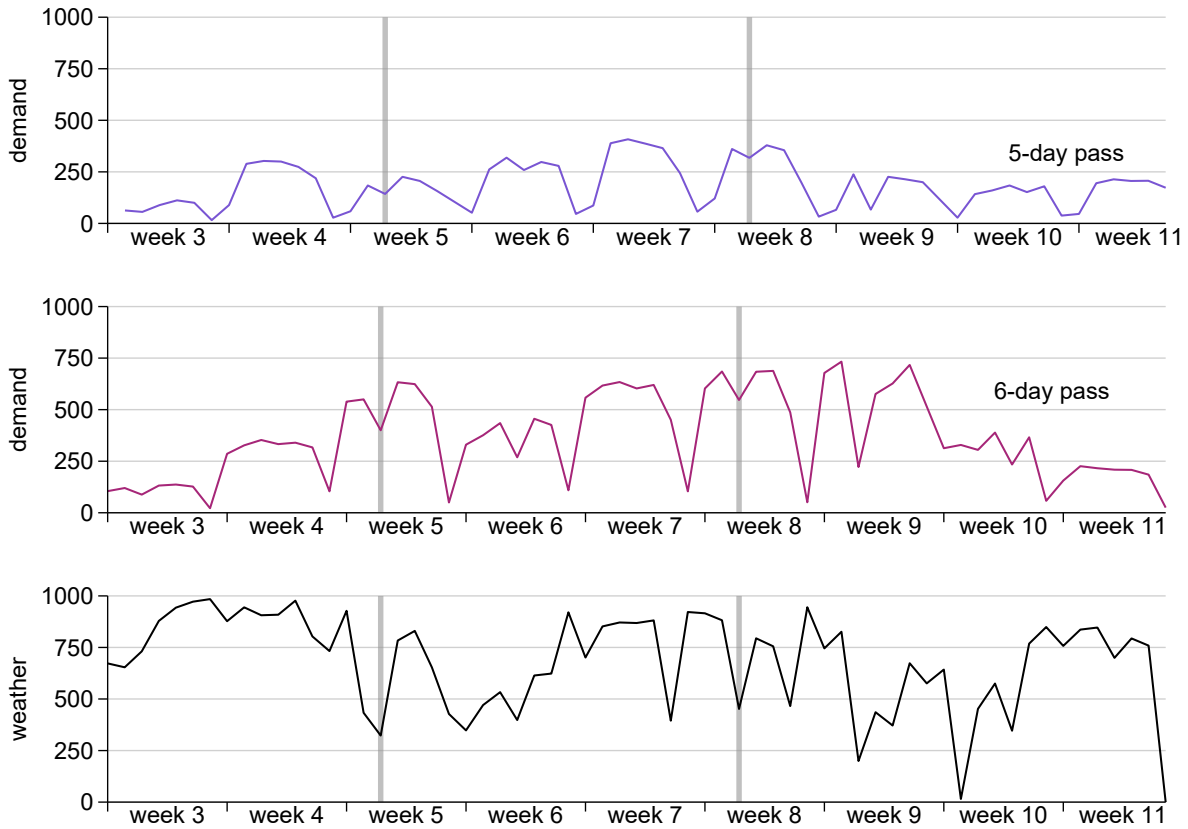
I estimate (11) by a least-squares dummy variable estimator (LSDV). The results are presented in Table 2. In all areas and over all forecast horizons the weather index has a larger impact on demand than the forecast error. Nevertheless, all coefficients are statistically significant. This suggests that **E**- and **L**-types appear to ski in all areas and confirming hypothesis (i). A change in the weather index by one standard deviation is associated with a 56%<sup>19</sup>, 51% and 83% change in skier demand in areas 1 to 3 holding all else equal, respectively. This indicates that in area 3 either more skiers base their decision upon the weather and/or forecast (higher  $\frac{N^E + N^E}{N}$ ) or that their elasticity with respect to the weather is higher (higher  $\bar{\gamma}$ ). The latter might be the case because of differences in guest structure. As mentioned in Section 3.1 the skiers buying one-day passes are quite heterogeneous regarding age and, potentially, regarding unobservable characteristics. The heterogeneity across age- and ticket-groups is evaluated later in this section. Table 2 indicates further the effects of forecast errors. I estimate that a change of a 0-day forecast error by one standard deviation changes demand, all else equal, significantly by 11.8%, 19.1% and 9.9% in area 1,2 and 3, respectively. Using the 0-day forecast I estimate the share of early deciders ( $N^E$ ) over the weather sensitive types ( $N^L + N^E$ ) in one-day passes being 38%, 70% and 28% for the three areas, respectively. Note that the large differences in coefficients between forecast errors and the weather stem not from fewer early deciders, but from fewer unstable days where the early deciders actually have an impact on the aggregate. Again, early deciders in stable weather situations (where the forecast error is null) are not visible as their decision is captured by the weather effect only.

---

<sup>19</sup>The Taylor approximation to interpret semi-elasticities in log-linear models as percentage changes is not feasible for large coefficients. Exact values are used here, where  $\% \Delta y = \exp(\Delta \beta) - 1$



Figure 4: Demand of 5-day pass and 6-day pass owners and weather in area 3, season 16/17.



**Figure Notes:** On the horizontal axis are the weeks between mid-January and mid-March in season 16/17 depicted. The vertical axis indicates daily demand for 5-day (top-panel) and 6-day passes (mid-panel) and the weather index scaled between 0 and 1000 (bottom-panel) all in area 3. The weather index incorporates the revealed preference variables of precipitation, sunshine duration and minimum temperature (see Section 3.5). Note that demand in 5-day passes is much higher from Mondays to Fridays whereas demand in 6-day passes is high from Sundays to Fridays as most owners of such passes spend their vacation from weekend to weekend. Two Tuesdays with relatively bad weather are highlighted in gray indicating that demand might take a weather-related drop even when most individuals have valid passes that day.

Figure 5: Two measures of model fit across areas and forecast horizons.

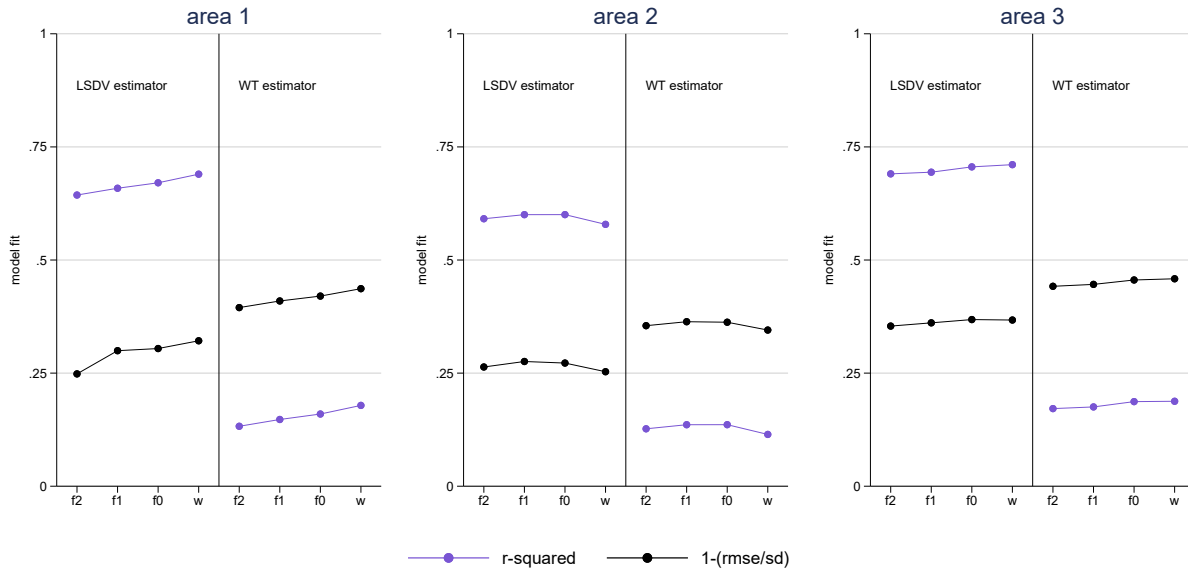


Figure Notes: The three panels indicate the model fit (vertical axis) for different forecast horizons ( $f^{-2}$ ,  $f^{-1}$ ,  $f^0$ ) and the weather  $w$  across two estimators using one-day passes in area 1, 2 and 3, respectively. The purple line is the in-sample fit  $R^2$  of (10) using level demand and the black line the out-of-sample fit proxy  $1 - \frac{RMSE}{SD}$  (one minus the root-mean-squared error divided by the standard deviation of demand) recovered from 10-fold cross validation.

Figure 6: Variable importance measure of RF including all available predictors

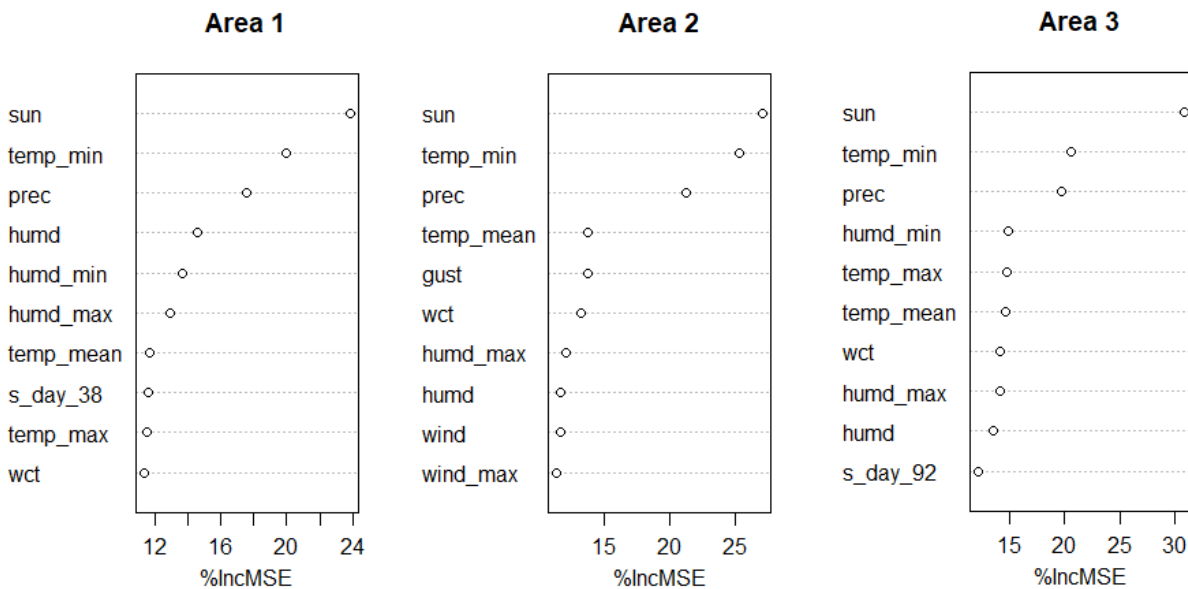


Figure Notes: Variable importance estimated by RF for each area. The measure on the x-axis indicates by how much the MSE would increase leaving the variable on the y-axis out of the random forest.

**Table 2: Effect of weather and forecast error on log demand**

variables	area 1			area 2			area 3		
	0-day forecast	1-day forecast	2-day forecast	0-day forecast	1-day forecast	2-day forecast	0-day forecast	1-day forecast	2-day forecast
weather ( $w$ )	0.025*** (0.0012)	0.025*** (0.0012)	0.026*** (0.0013)	0.020*** (0.0015)	0.021*** (0.0015)	0.021*** (0.0015)	0.030*** (0.0011)	0.030*** (0.0011)	0.030*** (0.0011)
forecast error ( $\Delta$ )	0.0093*** (0.0017)	0.0063*** (0.0017)	0.0043** (0.0015)	0.014*** (0.0019)	0.013*** (0.0018)	0.011*** (0.0017)	0.0083*** (0.0019)	0.0089*** (0.0018)	0.0068*** (0.0016)
easter ( $e$ )	1.01*** (0.19)	0.97*** (0.19)	0.97*** (0.20)	1.46*** (0.20)	1.41*** (0.19)	1.43*** (0.20)	0.89*** (0.24)	0.84*** (0.23)	0.87*** (0.24)
season day FE	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
$N$	910	910	910	1099	1099	1099	1154	1154	1154
$R^2$	0.787	0.783	0.780	0.674	0.671	0.667	0.789	0.791	0.788

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ 

**Table Notes:** The table depicts LSDV estimates of model (11) for three areas and three forecast horizons. Standard errors are clustered on the season day to account for intra-day correlations across seasons (seasonality). To allow a comparison between the areas only one-day passes are used. The weather ( $w$ ) and forecast ( $f^{-h}$ ) indices are continuous variables scaled between 0 and 100. The forecast error variable  $\Delta^{-h} = f^{-h} - w$  is the difference between weather and forecast. Easter ( $e$ ) is a dummy indicating the four Easter holidays (Good Friday to Easter Monday). All models are estimated using season day fixed effects to account for the seasonality.

In order to further investigate upon possible heterogeneities in guest structure and to elaborate on hypothesis (ii), I estimate (11) in area 2 and 3 (data in area 1 is only for one-day passes plausible) for three age groups (adults, juveniles, kids) and five pass-validity groups (one-day pass, weekend pass, one-week pass, two-week pass and season pass). Note that the latter four groups are aggregates for two- to four-day passes, five- to seven-day passes, eight- to fourteen-day passes and all else above fourteen-day passes. The results are depicted in Figure 7. Both areas have in common that the weather effect estimates in one-day and season pass groups are the largest. Indicating that in these groups there is a high average weather elasticity  $\bar{\gamma}$  or there are relatively more weather and forecast dependent skiers in these groups. This comes at no surprise: One-day pass owners consist of day-trippers with rather high transaction costs or weather sensitive overnight guests that decided against overtaking the weather risk from the area by buying multiple-day passes. Season pass owners, conversely, have low transaction cost but due to their proximity simply choose among the days with the best skiing conditions. In between these two extremes are two- to fourteen-day passes: Owners of such passes have rather high sunk costs. They overtake the weather risk from the ski area and try to make as much as possible out of their expenses. On top of that, parents might send their children and juveniles to a local ski school that teaches lessons in all conditions. This could also be the reason why the point estimates in weather effects for children and juveniles within all pass-validity groups are lower compared to adults. Apart from ski schools and ski clubs, most children and juveniles are constraint by the decision of their parents and would, by that, lead to the same estimates as for the adults.

The estimates of the forecast errors are in line with hypothesis (ii). Let us focus on one-day pass and season-pass owners only<sup>20</sup> in area 3. Although the same conclusions can be drawn from both areas, the differences in forecast error estimates are more pronounced in area 3. First, the share of early deciders among the weather sensitive is for the adults with one-day passes around 30% (The share  $\hat{\delta}/\hat{\beta}$  is visible in Figure 7 by dividing the point estimate of the forecast error by the point estimate of the weather). This stands in contrast to the season pass owners where the forecast error estimates are indistinguishable from zero and, consequently, consist of late deciders almost exclusively. Because one-day pass owners have arguably much higher transaction costs than season pass owners, the data supports hypothesis (II). Secondly, among one-day pass owners the juveniles and children have clearly the highest estimated share of early deciders. This observation also supports hypothesis (II) because within the one-day pass owners families have clearly the highest transaction costs.

---

<sup>20</sup>Note that all passes between two- to fourteen days indicate rather conflicting results. As the weather and forecast dependence is generally low and confidence intervals often surpass zero in these groups, it allows no meaningful interpretation of forecast error estimates and early deciders share at this aggregation level.

Hypothesis (iii) is tested by including a slope dummy in (11) for optimistic forecasts. I estimate

$$\log(y_{ds}) = w_{ds}\beta + \Delta_{ds}\delta + \Delta_{ds} \cdot D_{ds}^o \nu + \alpha_d + e_{ds}\gamma + \varepsilon_{ds}, \quad (18)$$

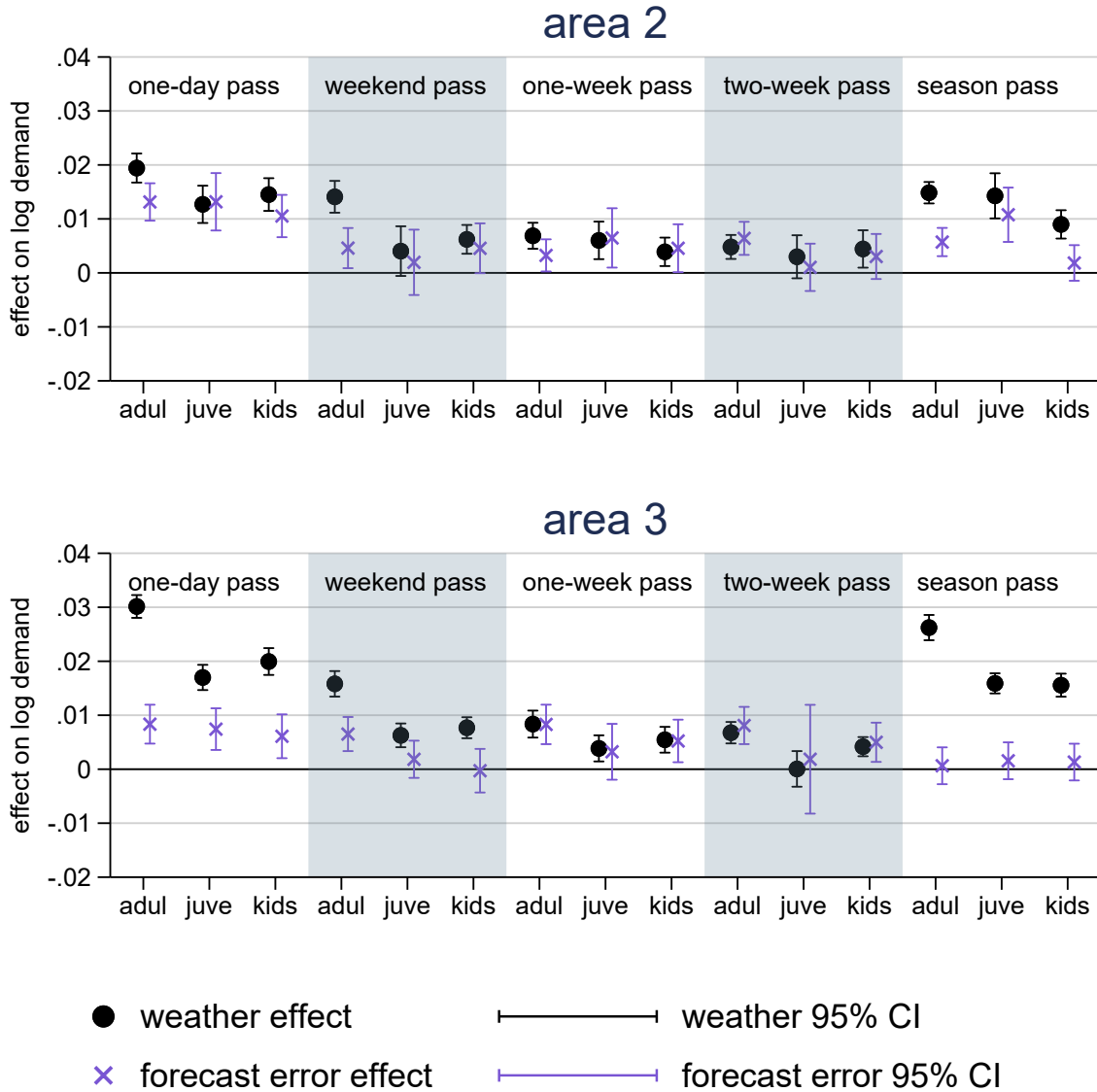
where  $D_{ds}^o = \mathbf{1}[(f_{ds}^{d-h} - w_{ds}) > 0]$ <sup>22</sup> is a dummy indicating an optimistic forecast and the other variables are the same as in (11). With this specification I test whether an optimistic forecast affects demand different than a pessimistic forecast. The LSDV estimates are depicted in Table 3. The semi-elasticities for pessimistic and optimistic forecasts are recovered by  $\frac{\partial \log(y_{ds})}{\partial \Delta_{ds}} = \delta + D^o \cdot \nu$  with  $D^o = 0$  and  $D^o = 1$ , respectively. Pessimistic forecast errors are statistically significant over almost all areas and forecast horizons. For example, I estimate that a pessimistic 0-day forecast error of one standard deviation reduces demand in area 1 *c.p.* by 21% compared to a day with an accurate forecast. On the contrary, optimistic forecast errors seem to affect skier demand only in area 2 and slightly but less clear in area 3. An optimistic 0-day forecast error of one standard deviation increases demand by 16% compared to an accurate weather forecast in area 2. Area 2 has also the highest share of early deciders among one-day pass owners and strong impacts of errors in both directions. These might be consequences of its famous family friendliness and their high transaction costs. But more noticeable, throughout all models the effect of optimistic forecasts are smaller than those of pessimistic forecasts and, by that, supporting hypothesis (iii).

---

<sup>21</sup> Anyone that went on a day trip with her family as a child or as adult remembers the struggle of getting up early in the morning, dressing the children, packing all the gear into the car and walking from the parking spot to the area entrance. Moving a whole family to a ski area for a day-trip requires much more effort than a single person. Sticking to a clear plan and preparing the gear on the eve before the day simplifies matters a lot. Therefore, it would not be surprising if most families decided early on.

<sup>22</sup> The forecast error  $\Delta_{ds}$  is here demeaned in each area to be centred around zero. As the spatial matching of weather and forecast data is an approximation (see Section 3.2), the forecast indices are slightly lower than the weather indices and render forecasts more pessimistic than they actually are.

Figure 7: Effects of weather and forecast errors on log demand separated by age and pass-validity groups



**Figure Notes:** The figure indicates the point estimates of (11) for weather and forecast errors on log demand in area 2 (upper panel) and area 3 (lower panel). Along the x-axis for each of the five pass-validity groups are three age groups. The pass-validity groups are from left to right aggregates for one-day passes, two- to three-day passes, four- to seven-day passes, eight- to fourteen-day passes and all else above fourteen-day passes. The age-groups are denoted as “adult” for adults, “juve” for juveniles and “kids” for children. Points indicate point estimates and the bandwidths 95% confidence intervals of the point estimates. Standard errors are clustered on the season day to account for intra-day correlations across seasons (seasonality).

**Table 3: Effect of weather and forecast error with slope dummy on log demand**

variables	area 1			area 2			area 3		
	0-day forecast	1-day forecast	2-day forecast	0-day forecast	1-day forecast	2-day forecast	0-day forecast	1-day forecast	2-day forecast
weather ( $w$ )	0.023*** (0.0013)	0.025*** (0.0014)	0.025*** (0.0015)	0.020*** (0.0015)	0.021*** (0.0015)	0.021*** (0.0015)	0.029*** (0.0012)	0.030*** (0.0012)	0.030*** (0.0012)
fc.e. pessimistic ( $\Delta D^o = 0$ )	0.016*** (0.0033)	0.0078* (0.0036)	0.0050 (0.0029)	0.017*** (0.0040)	0.013** (0.0043)	0.0099* (0.0039)	0.0098** (0.0036)	0.011** (0.0033)	0.0063* (0.0030)
fc.e. optimistic ( $\Delta D^o = 1$ )	0.000 (0.004)	0.004 (0.004)	0.003 (0.004)	0.012*** (0.003)	0.013** (0.004)	0.012*** (0.003)	0.007* (0.003)	0.007* (0.003)	0.007** (0.003)
easter ( $e$ )	1.02*** (0.18)	0.97*** (0.19)	0.98*** (0.20)	1.46*** (0.20)	1.41*** (0.19)	1.43*** (0.20)	0.90*** (0.24)	0.85*** (0.24)	0.87*** (0.24)
season day FE	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
$N$	910	910	910	1099	1099	1099	1154	1154	1154
$R^2$	0.790	0.783	0.780	0.674	0.671	0.667	0.789	0.791	0.788

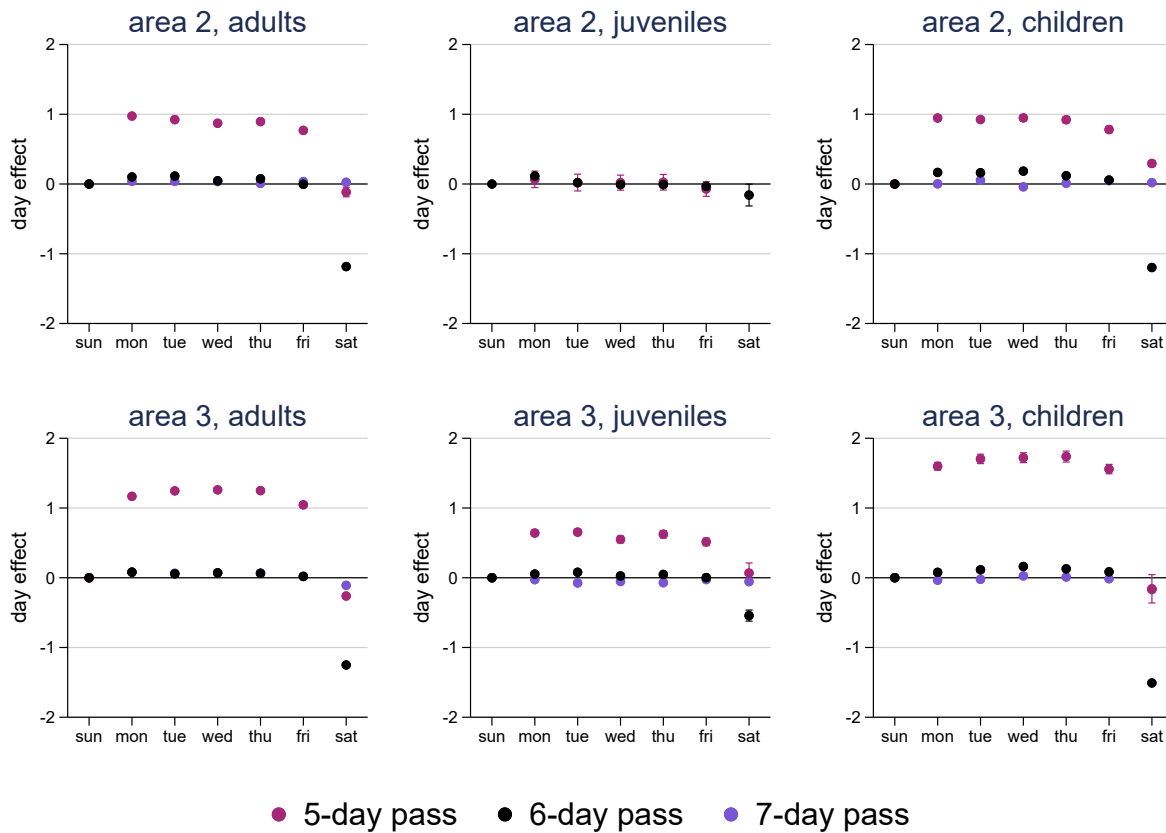
Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ 

**Table Notes:** The table depicts LSDV estimates of model (18) for three areas and three forecast horizons. Standard errors are clustered on the season day to account for intra-day correlations across seasons (seasonality). To allow a comparison between the areas only one-day passes are used. The weather ( $w$ ) and forecast ( $f^{-h}$ ) indices are continuous variables scaled between 0 and 100. The forecast error variable  $\Delta^{-h} = f^{-h} - w$  is the difference between weather and forecast. The dummy variable  $D^o = \mathbb{1}[(f^{-h} - w) > 0]$  is interacted with the forecast error variable  $\Delta^{-h}$  to allow for a slope change in optimistic forecasts. Hence, the semi-elasticities in the forecast errors are recovered by  $\frac{\partial \log(y_{ds})}{\partial \Delta_{ds}} = \delta + D^o \cdot \nu$  with  $D^o = 0$  and  $D^o = 1$ , respectively. Easter ( $e$ ) is a dummy indicating the four Easter holidays (Good Friday to Easter Monday). All models are estimated using season day fixed effects to account for the seasonality.

The results of estimating (17) by an LSDV estimator are depicted in Figure 8. Apart from the obvious demand drops on the weekends in 5-day and 6-day pass owners, there is no pattern visible. Demand of adult 5-day pass owners in area 2 is estimated to be twice as high on Mondays compared to Sundays in a comparable week controlling for weather. A slightly smaller coefficient results on Fridays (0.77 versus 0.97). Therefore skiers marginal utility is on average not decreasing in repeated consumption and, by that, Hypothesis (I) cannot be confirmed. Note that this only holds on average. It might still be true that individuals have either decreasing or increasing marginal utilities regarding repeated consumption. There is just no way to recover these from my data.

**Figure 8: Point estimates of week-day effects for three different passes in two areas over three age groups.**



**Figure Notes:** All panels indicate point estimates of week-day effects ( $\delta_{\bar{d}}$ ) in model (17) with the Sundays as reference categories. The three columns separate the estimates by age groups and the two rows by areas 2 and 3. The vertical bars indicate 95% confidence intervals of the estimates but remain mostly invisible due to very small bands. 7-day pass owners indicate no preference across all areas and age groups. 6-day pass owners clearly ski less on Saturdays due to mostly arriving or departing that day and 5-day pass owners show a preference for non-weekends.

To evaluate Hypothesis (II) the weather index is interacted with the week-day dummies



to test whether the weather dependence changes throughout a week. The modified model reads then

$$\log(y_{\tilde{d}ks}) = \beta w_{\tilde{d}ks} + \delta' \tilde{d} + \gamma' \tilde{d} \cdot w_{\tilde{d}ks} + \alpha_k + e_{\tilde{d}ks} + \varepsilon_{\tilde{d}ks}, \quad (19)$$

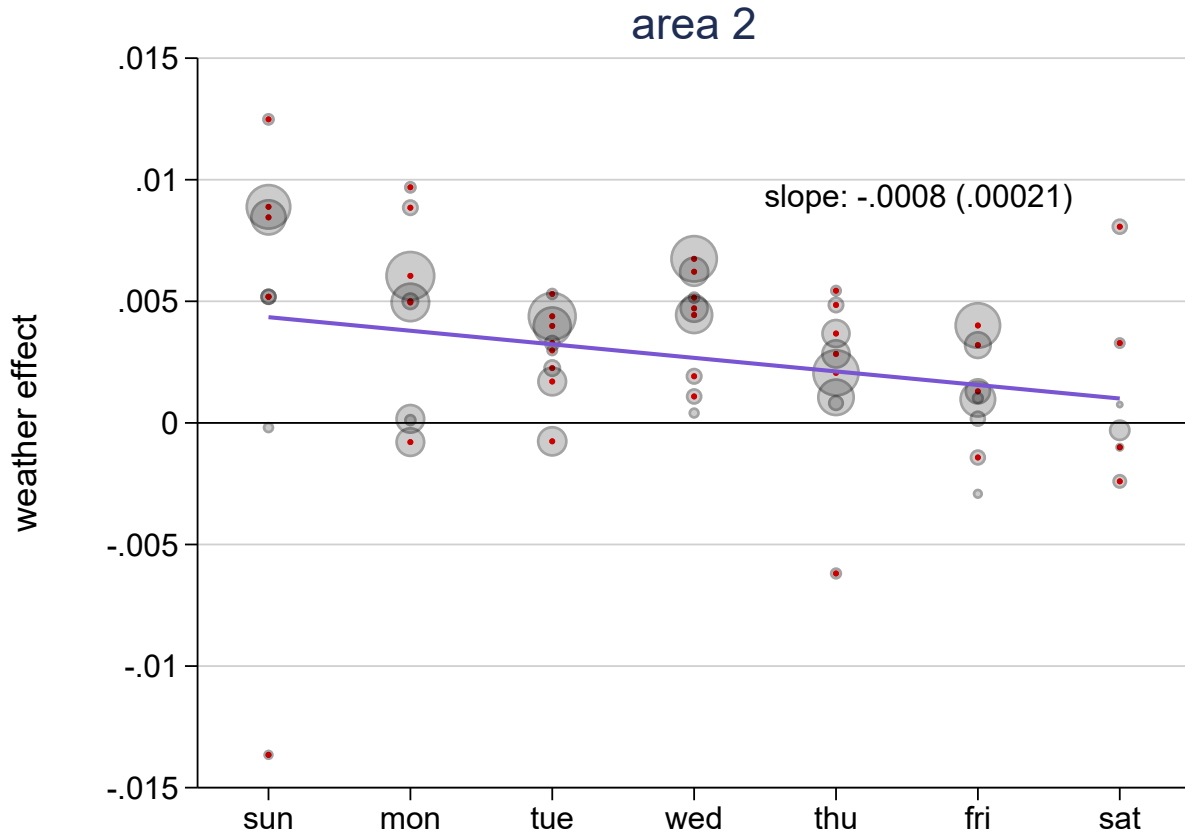
where all variables and coefficients remain as in (17) and  $\gamma$  indicates the vector of coefficients of the week-day and weather interaction. Essentially, (19) exploits the within week-day variation of one-week pass owners and estimates the weather elasticities separately for each week-day. As I recover the partial effect of the weather  $\frac{\partial \log(y_{\tilde{d}ks})}{\partial w_{\tilde{d}ks}} = \hat{\beta} + \hat{\gamma}' \cdot \tilde{d}$  for each of the 7 week-days and this model can be estimated 9 times (across three age groups and three pass-validity groups), I receive 62 point estimates for each area. Lining those up on the weekdays and apply a demand-weighting (the summed up demand upon which each estimate is drawn) allows me to estimate the correlation of weather effects and week-days by a weighted least-squares (WLS) estimator. The results are depicted in Figures 9 and 10. I find that the weather effect substantially decreases to the end of the week and by that support Hypothesis (II). The two figures suggest, that the weather effect decreases by 0.0008 and 0.01 per day for area 2 and 3, respectively. To put this in perspective in area 3: When on Sundays a decrease in the weather index by 10 (approximately one half standard deviation of the weather index) hold on average around 7.5% individuals off skiing, the same decrease in the weather index on Fridays reduces demand only by roughly 2.5%. Hence demand is affected by roughly one percentage point less per day of the week from the weather. Because of the infeasibility of disentangling  $\frac{N^W}{N}$  and  $\bar{\gamma}$  in (16) it remains unclear whether the individual's weather elasticity is decreasing over the span of a week or just fewer individuals decide upon the weather. Nevertheless, both show the circumstance that week-pass owners on average become less sensitive to the weather at the end of their vacation.

To evaluate Hypothesis (III) the forecast index is interacted with the week-day dummies to test whether the forecast error dependence changes throughout a week. The modified model reads then

$$\log(y_{\tilde{d}ks}) = \beta w_{\tilde{d}ks} + \delta' \tilde{d} + \nu \Delta_{\tilde{d}ks} + \gamma' \tilde{d} \cdot \Delta_{\tilde{d}ks} + \alpha_k + e_{\tilde{d}ks} + \varepsilon_{\tilde{d}ks}, \quad (20)$$

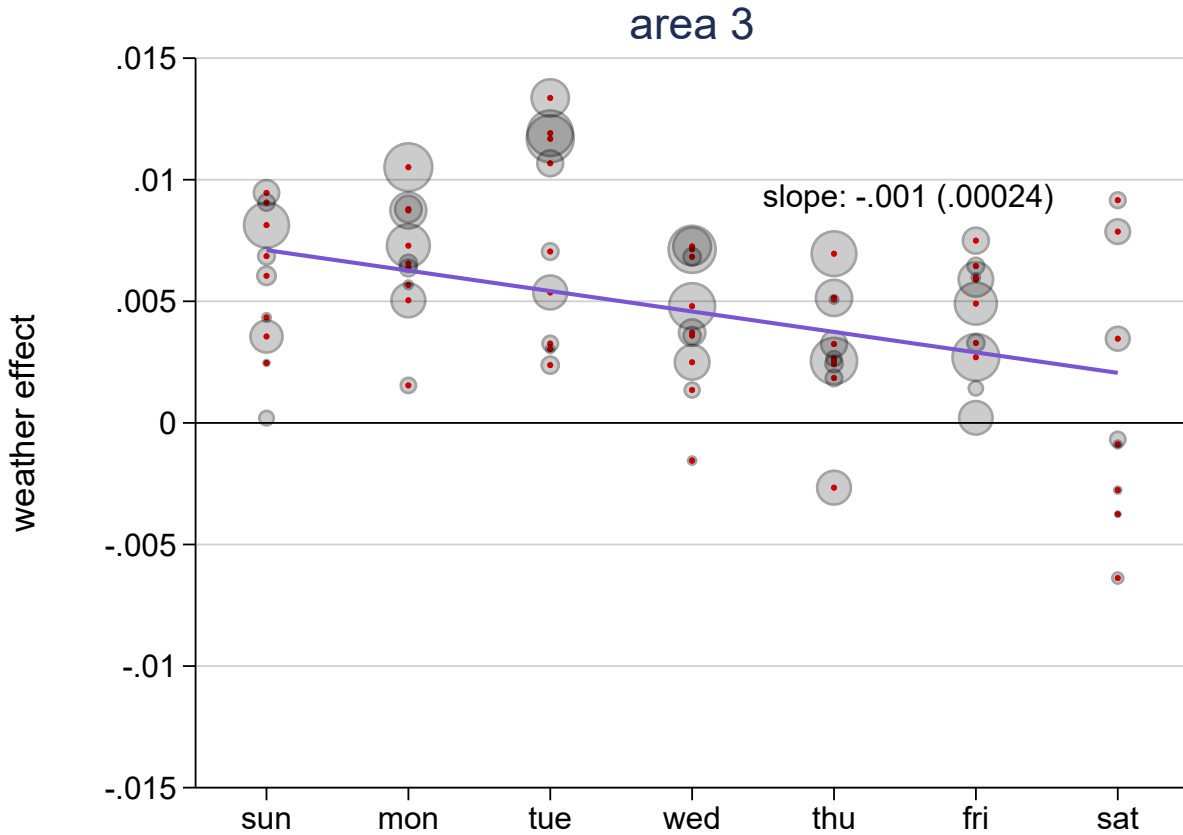
where all variables and coefficients remain as in (19) except that  $\nu$  is the coefficient on the forecast error and  $\gamma$  indicates the vector of coefficients of the week-day and forecast error interaction. Essentially, (20) exploits the within week-day variation of one-week pass owners and estimates the forecast error elasticities separately for each week-day while controlling for the weather. This is necessary as a forecast error is a proxy of the forecast deviation from the weather and as such can only deviate in the direction given a certain weather. In particular, a forecast error cannot be negative when the weather is already at its worst. As I

Figure 9: Point estimates of partial weather effects in area 2 across week-pass groups and three age groups.



**Figure Notes:** All points indicate single point estimates of partial weather effects  $(\hat{\beta} + \hat{\gamma}' \cdot \tilde{d})$  in model (19) weighted by the overall demand of the days involved in estimating (19). The sample consists of week-pass owners (five- to seven-day passes). Standard errors are clustered on the week-level ( $k$ ) to account for intra-cluster correlations. Statistical significance at the 5%-level is indicated by red dots within the points. Larger points indicate an estimate that is backed by more demand and receives a relatively high weighting in the WLS estimation. The purple line gives the linear fit of the WLS estimation and results in a negative correlation between weather effects and week-days. The slope coefficient of this correlation and its standard error are provided directly in the graph.

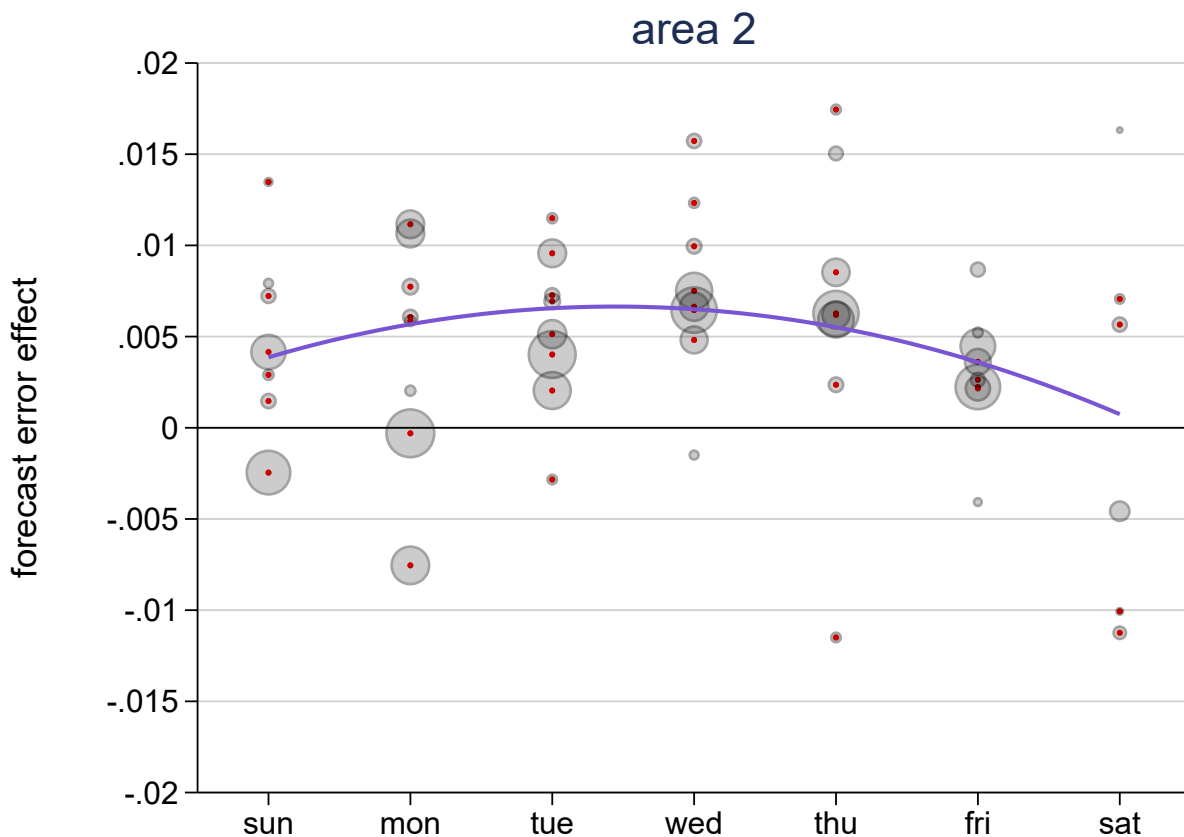
Figure 10: Point estimates of partial weather effects in area 3 across week-pass groups and three age groups.



**Figure Notes:** All points indicate single point estimates of partial weather effects  $(\hat{\beta} + \hat{\gamma}' \cdot \tilde{d})$  in model (19) weighted by the overall demand of the days involved in estimating (19). The sample consists of week-pass owners (five- to seven-day passes). Standard errors are clustered on the week-level ( $k$ ) to account for intra-cluster correlations. Statistical significance at the 5%-level is indicated by red dots within the points. Larger points indicate an estimate that is backed by more demand and receives a relatively high weighting in the WLS estimation. The purple line gives the linear fit of the WLS estimation and results in a negative correlation between weather effects and week-days. The slope coefficient of this correlation and its standard error are provided directly in the graph.

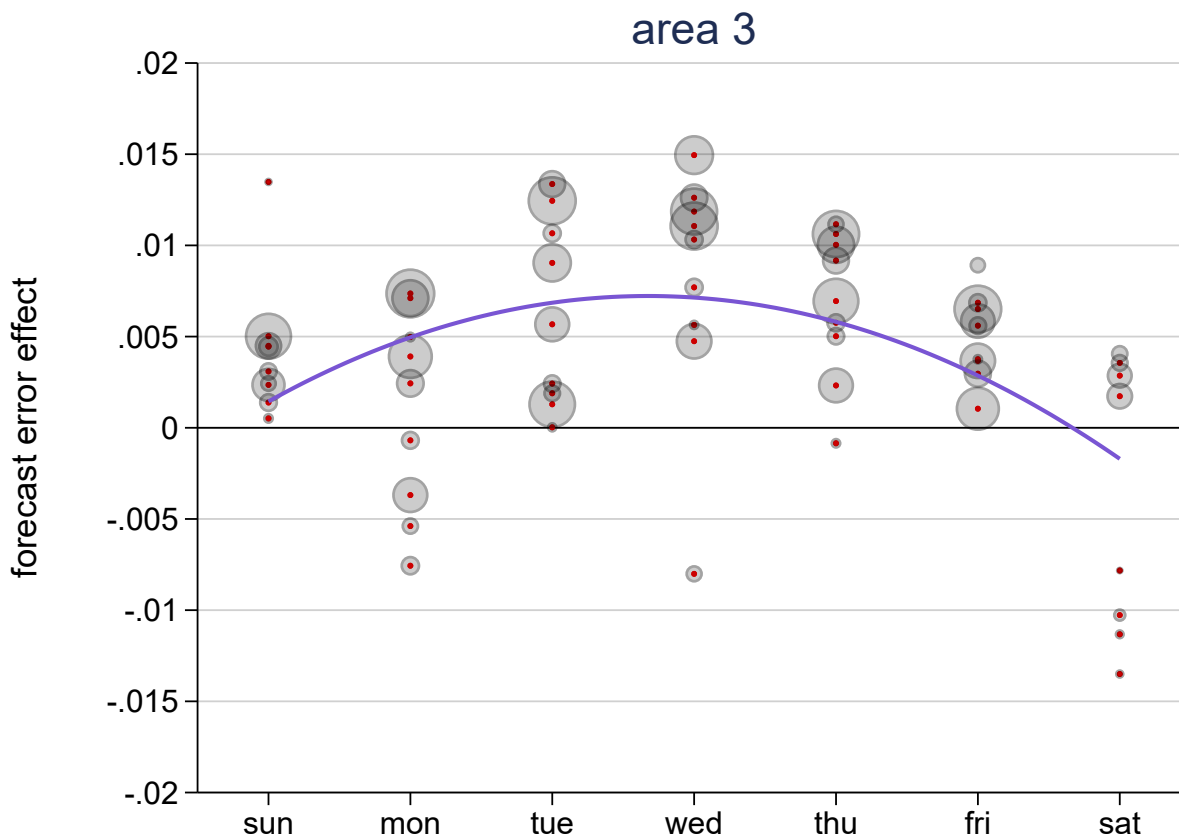
recover the partial effect of the forecast error  $\frac{\partial \log(y_{dks})}{\partial \Delta_{\tilde{d}_{ks}}} = \hat{\nu} + \hat{\gamma}' \cdot \tilde{d}$  for each of the 7 week-days and this model can be estimated 9 times (across three age groups and three pass-validity groups), I receive 62 point estimates for each area. Lining those up on the weekdays and apply a demand-weighting (the summed up demand upon which each estimate is drawn) allows me to estimate the correlation of weather effects and week-days by a weighted least-squares (WLS) estimator. The results are depicted in Figures 11 and 12. Clearly, there seems to be a quadratic relationship between forecast error effects and days of the week in both areas. This is in line with Hypothesis (II) and (III): The role of the weather and the forecast are diminished as probably most week-pass owners want to make full use of their last day in vacation, regardless of the weather. But at the start of the week things look different. Most individuals decide rather late but still weather sensitive. In that respect, the forecast is not of much use. Whereas at mid-week most season-pass owners seem to decide early upon forecasts. The reason for this might be quite obvious: It is the perfect moment to take a break from skiing and engage in other activities to recover strength for the last two days.

Figure 11: Point estimates of partial forecast error effects in area 2 across week-pass groups and three age groups.



**Figure Notes:** All points indicate single point estimates of partial forecast error effects ( $\hat{\nu} + \hat{\gamma}' \cdot \tilde{d}$ ) in model (20) weighted by the overall demand of the days involved in estimating (20). The sample consists of week-pass owners (five- to seven-day passes). Standard errors are clustered on the week-level ( $k$ ) to account for intra-cluster correlations. Statistical significance at the 5%-level is indicated by red dots within the points. Larger points indicate an estimate that is backed by more demand and receives a relatively high weighting in the WLS estimation. The purple line gives the quadratic fit of the WLS estimation.

Figure 12: Point estimates of partial forecast error effects in area 3 across week-pass groups and three age groups.



**Figure Notes:** All points indicate single point estimates of partial forecast error effects ( $\hat{\nu} + \hat{\gamma}' \cdot \tilde{d}$ ) in model (20) weighted by the overall demand of the days involved in estimating (20). The sample consists of week-pass owners (five- to seven-day passes). Standard errors are clustered on the week-level ( $k$ ) to account for intra-cluster correlations. Statistical significance at the 5%-level is indicated by red dots within the points. Larger points indicate an estimate that is backed by more demand and receives a relatively high weighting in the WLS estimation. The purple line gives the quadratic fit of the WLS estimation.

## 6 Conclusion

The impact of the weather and its forecast on ski demand is strongest in one-day pass owners. In a static setting potential one-day pass owners form their decision regarding their activity choice day-by-day. Using this framework I show that a substantial share of aggregate demand base their choice on the expectation of weather outcomes. It appears that between 20 and 25% of the variation in demand can be explained by weather outcomes alone. As skiing is a typical leisure activity that involves high transaction cost, the seasonality of the weekends, Christmas, Easter and school holidays is determined by an individual daily activity constraint. This constraint amounts to around 50% of the variation in the aggregate. Furthermore, the daily weather is seasonal too and is largely affected by large scale situations that are more or less persistent over time. The weather service's forecast accuracy varies with these situations and impacts the risk evaluation of the weather sensitive skier. In particular, the less stable the weather situation, the more risky an early decision based on forecasts become. From this it follows that potential skiers are incentivized to postpone their skiing decision as close to the activity as possible. But, as potential late deciders face increasing transaction costs, some opt for early decisions and might fail in their individual optimization when forecasts deviate from the weather. In the aggregate, I estimate that forecast errors of one standard deviation amount to changes in demand up to 20%. Observing variation in weather and forecast errors allows me to estimate the share of early deciders among the weather sensitive skiers. This share amounts to 38%, 70% and 28% in area 1, 2 and 3, respectively.

Separating demand into different age groups and pass-validity groups in area 2 and 3 reveals that those groups related to high transaction costs consist of a larger share of early deciders. The best way to see this is comparing one-day pass owners to season pass owners in area 3. The former face arguably relatively large transaction costs compared to the latter. I estimate that 30% of the one-day pass owners consist of early deciders whereas all of the season-pass owners consist of late deciders. On top of that, children and juveniles, often accompanied by their families, reveal of being consistently more affected by forecasts and consist, thus, of higher shares of early deciders. This is in line with the idea that high transaction costs enforce an early decision.

By introducing the option for individuals to choose among a third weather dependent but much cheaper alternative activity, I suggest that early deciders face stronger losses in potential optimistic forecasts compared to potential pessimistic forecasts. The losses from a suboptimal early decision against skiing, due to better than expected weather, can be mitigated by following the alternative outdoor activity at home. On the contrary, the

losses from a suboptimal early decision in favour of skiing, due to worse than expected weather, cannot be mitigated as transaction costs have been born already and these sunk costs incentivize the skier to enter the area in unfortunate weather. Including a slope dummy on the forecast error supports this hypothesis as all estimates in pessimistic forecasts have larger coefficients than in the optimistic case. The emergence of more low-cost outdoor activities coupled with less fog days and milder winter weather might actually contribute to a large extent to the stagnating skiing demand within the Alps (Plaz & Schmid, 2015).

From these learnings I suggest that ski area operators should pay attention to unstable weather situations when implementing new pricing regimes. In particular, prices should be conditional on forecast uncertainty rather than on weather forecasts itself. The loss invoked on an individual through an optimistic forecast can be large already. Extracting more out-of-pocket expenses in such a situation could alienate the affected further. On the contrary, lowering the price for a day with a potential optimistic forecast could mitigate losses and lure additional skiers to the area that would otherwise rest at home. Or, instead of using dynamic prices, an operator could refund sold tickets in the case of worse weather outcomes and allow them to switch the price of ski passes for alternative activities within the destination. In practice, such a refund works better when firms are horizontally integrated within the area. Similar to the reduced prices, this could dampen the losses from an optimistic forecast.

In the last part I look at the dynamic choices of week-pass owners and find that they do not fully make use of their week-pass by skipping some days with unfortunate weather. The estimates indicate that weather sensitivity decreases to the end of the week suggesting that week-pass owners make the most out of their last days of vacation or that they become polarized in their motivation. Namely that they either decrease or increase their skiing engagement due to deteriorating endurance or increasing skills. The decision to ski at the last days hinges then less on the weather but to a larger extent on these motivational factors. Looking at forecast errors reveals that mid-week days encounter the largest drop in demand when forecasts are unfortunate. The reason for this might be that skiers tend to use these days for a break to regain strength for the last days and more often, when the weather forecast for such a day is unfortunate. As the de-motivated skiers form a regret on their already paid week-passes to the end of the week, they might fail in their optimization and ski in harsh conditions. Introducing refundable days in week-passes would help these individuals to overcome their regrets and follow their optimal activity on such a day.



## 7 Literature

### References

- Malasevska, I., Haugom, E., & Lien, G. (2017a). Optimal weather discounts for alpine ski passes. *Journal of Outdoor Recreation and Tourism*, 20(May), 19–30.
- Shih, C., Nicholls, S., & Holecek, D. F. (2009). Impact of Weather on Downhill Ski Lift Ticket Sales. *Journal of Travel Research*, 47(3), 359–372.
- Malasevska, I., & Haugom, E. (2018). Optimal prices for alpine ski passes. *Tourism Management*, 64, 291–302.
- Koenig, U., & Abegg, B. (1997). Impacts of climate change on winter tourism in the swiss alps. *Journal of Sustainable Tourism*, 5(1), 46–58.
- Elsasser, H., & Bürki, R. (2002). Climate change as a threat to tourism in the Alps. *Climate Research*, 20(3), 253–257.
- Gössling, S., Scott, D., Hall, C. M., Ceron, J. P., & Dubois, G. (2012). Consumer behaviour and demand response of tourists to climate change. *Annals of Tourism Research*, 39(1), 36–58.
- Gonseth, C. (2013). Impact of snow variability on the Swiss winter tourism sector: Implications in an era of climate change. *Climatic Change*, 119(2), 307–320.
- Steiger, R., & Abegg, B. (2017). Aktuelle und zukünftige Schneesicherheit der Skigebiete in Graubünden. (January 2015).
- Steiger, R., & Abegg, B. (2018). Ski areas' competitiveness in the light of climate change: Comparative analysis in the Eastern Alps. *Tourism in Transitions*, 187–199.
- Plaz, P., & Schmid, S. (2015). Aussichten für das Geschäft mit den alpinen Schneesportlern in Graubünden. Grundlagenberichtim Rahmen des Projekts „Strategien für Bündner Tourismusorte“.
- Falk, M. (2015). Return on Investments in New Ski Lifts : The Importance of Weather Conditions and Elevation. (February).
- Malasevska, I., Haugom, E., Hinterhuber, A., Lien, G., & Mydland, Ø. (2020). Dynamic pricing assuming demand shifting: the alpine skiing industry. *Journal of Travel and Tourism Marketing*, 37(7), 785–803.
- Lütolf, P., Stettler, J., Wagenseil, U., Wallimann, H., & Rohrer, B. (2020). Auswirkungen der neuen Preismodelle - Erkenntnisse für die Bergbahnbranche. *SBS Schriften*, 10, 241.
- Scaglione, M., & Doctor, M. (2008). The impact of inaccurate weather forecasts on cable-car use, 1–12.

- MeteoSchweiz. (2012). Dem Wetter einen Schritt voraus: Von der Messung bis zur Prognose, 24 pp.
- Diamond, R., McQuade, T., & Qian, F. (2017). The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco. *NBER Working Paper*.
- Scott, P. T. (2014). Dynamic Discrete Choice Estimation of Agricultural Land Use. (May).
- Katz, R. W., Murphy, A. H., & Winkler, R. L. (1982). Assessing the Value of Frost Forecasts to Orchardists: A Dynamic Decision-Making Approach. *Journal of Applied Meteorology*, 21.
- MeteoSchweiz. (2015). *Typische Wetterlagen im Alpenraum* (tech. rep.).
- Rutty, M., & Andrey, J. (2014). Weather forecast use for winter recreation. *Weather, Climate, and Society*, 6(3), 293–306.
- Haugom, E., & Malasevska, I. (2019). The relative importance of ski resort- and weather-related characteristics when going alpine skiing. *Cogent Social Sciences*, 5(1).
- Osczevski, B. Y. R., & Bluestein, M. (2005). Equivalent Temperature. *Van Nostrand's Scientific Encyclopedia*, (May), 1453–1458.
- Malasevska, I., Haugom, E., & Lien, G. (2017b). Modelling and forecasting alpine skier visits. *Tourism Economics*, 23(3), 669–679.
- Holmgren, M. A., & McCracken, V. A. (2014). What Affects Demand for "The Greatest Snow On Earth?". *Journal of Hospitality Marketing and Management*, 23(1), 1–20.
- Hastie, T., Tibishirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning*.
- Lovell, M. C. (2008). A Simple Proof of the FWL Theorem. *The Journal of Economic Education*, 39(1), 88–91.
- Carson, R., & Yu, C. A. (2020). Sources and Nature of Measurement Error in Estimating Climate Related Impacts, 1–66.

## 8 Appendix

### 8.1 Spatial join of weather data to lift stations

To spatially join the highest and lowest lift with all weather stations within 30km, we use ArcGIS software to match geo-referenced data of our highest and lowest lifts and of all weather stations from MeteoSwiss<sup>23</sup>. We find two types of weather stations: Those only measuring precipitation and those measuring almost all variables. The distance weighting

---

<sup>23</sup>found here: <https://www.meteoschweiz.admin.ch/home/messwerte.html?param=messnetz-automatisch>

has to be calculated for each variable in each year to account for missing data in a given year and station. The basis of these calculations is drawn from (Carson & Yu, 2020) section 3.1.2:

$$\hat{w}_{it} = \frac{\frac{w_{A,t}}{d_{A,i}^p} + \frac{w_{B,t}}{d_{B,i}^p}}{\frac{1}{d_{A,i}^p} + \frac{1}{d_{B,i}^p}} \quad (21)$$

Where  $i$  is area/lift  $i$ ,  $t$  is the year,  $A$  and  $B$  are the two weather stations within the 30km radius.  $\hat{w}$  depicts the estimated weather variable,  $w$  the weather input variables and  $d$  the distance between stations  $A$ ,  $B$  and area/lift  $i$ .  $p$  is a power parameter that penalizes distance (usually  $p = 2$  according to (Carson & Yu, 2020)). Rearranged and evaluated for  $n$  weather stations in the vicinity of 30km we get the weight  $\omega$  for station  $A$  at time  $t$ :

$$\omega_{A,t} = \frac{\frac{1}{d_{A,i}^p}}{\sum_{s=A}^n \frac{1}{d_{s,i}^p}} \quad (22)$$

for each weather station. The final weather variable is then the sum over all weather variables weighted by stations:

$$\hat{w}_{it} = \sum_{s=A}^n w_{s,t} * \omega_{s,t} \quad (23)$$

## 8.2 Impact of Season Fixed Effect

To check upon the impact of a season fixed effect, I compare log-linear estimates (10) across the three areas. Point estimates are in Tables 4, 5 and 6. Column (1) and (2) indicate the models without and with the season fixed effect  $\varsigma_s$  for all three forecast horizons and the weather, respectively. The tables indicate the OLS estimates of the coefficients and standard errors in parentheses. These are clustered on the season day to account for the likely intra-cluster correlation of the seasonality. In all areas and over all forecast horizons the sunshine is significantly positive related to skiing demand. The opposite is true for measured precipitation: There is a significant negative relationship to demand but only up to the cubic term in area 1 and only in areas 1 and 3 additionally in the forecast specifications. Measured wind-chill temperature is significantly related to demand in area 1 and 2, the forecast wind-chill temperature in area 2 and 3. I refrain from the interpretation of any effect sizes due to the *ceteris paribus* property of OLS estimates. Despite the exogeneity of the regressors, in reality it is hardly ever the case that a single weather variable changes while the others remain constant.

Comparing the estimates in (1) and (2) reveals that the season fixed effect has in almost all cases no significant confounding effect on any of the regressors. The exception are the

measured wind-chill temperature effect sizes in area 3. After incorporating the season fixed effects, the effects are half the size and are no longer statistically significant. A possible explanation could be infrastructure investments in temperature and wind exposed lifts (e.g. chairlift with bubbles) that led demand to being less dependent on wind-chill temperature during the observed time period. However, with the goal of exploiting as much variation as possible, the season fixed effects are disregarded in most of the following analysis.

**Table 4: Log demand and its association to weather and forecast variables in area 1**

variables	Forecast d-2		Forecast d-1		Forecast d-0		Weather	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
fsnow	-0.00022 (0.0029)	-0.0011 (0.0028)	-0.0021 (0.0036)	-0.0026 (0.0034)	0.00054 (0.0051)	-0.00017 (0.0049)	-0.0020 (0.0060)	-0.0027 (0.0057)
prec	-0.12*** (0.025)	-0.12*** (0.025)	-0.13*** (0.031)	-0.13*** (0.029)	-0.11*** (0.024)	-0.12*** (0.023)	-0.16*** (0.027)	-0.16*** (0.026)
prec <sup>2</sup>	0.0046* (0.0018)	0.0042* (0.0018)	0.0063* (0.0028)	0.0063* (0.0026)	0.0050** (0.0015)	0.0053*** (0.0015)	0.0073*** (0.0021)	0.0073*** (0.0020)
prec <sup>3</sup>	-0.000068* (0.000031)	-0.000060 (0.000031)	-0.00011 (0.000060)	-0.00011 (0.000057)	-0.000071** (0.000023)	-0.000074*** (0.000022)	-0.00010** (0.000036)	-0.00010** (0.000035)
sun	0.0056*** (0.0011)	0.0057*** (0.0011)	0.0049*** (0.0011)	0.0049*** (0.0010)	0.0059*** (0.0011)	0.0057*** (0.0011)	0.0082*** (0.00072)	0.0082*** (0.00072)
wct	-0.021 (0.014)	-0.026 (0.015)	-0.020 (0.014)	-0.027 (0.015)	-0.016 (0.016)	-0.022 (0.016)	-0.022* (0.0091)	-0.022* (0.0094)
wct <sup>2</sup>	-0.0012 (0.00064)	-0.0015* (0.00068)	-0.0011 (0.00062)	-0.0015* (0.00067)	-0.00098 (0.00072)	-0.0013 (0.00074)	-0.0015*** (0.00042)	-0.0016*** (0.00043)
easter	1.07*** (0.20)	1.05*** (0.20)	0.96*** (0.22)	0.94*** (0.21)	0.96*** (0.23)	0.94*** (0.24)	1.10*** (0.19)	1.08*** (0.20)
season FE	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>
season day FE	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
<i>N</i>	910	910	910	910	910	910	910	910
<i>R</i> <sup>2</sup>	0.776	0.786	0.757	0.768	0.738	0.748	0.785	0.794

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Table Notes:** The table depicts LSDV estimates of model (10) for three forecast horizons and the weather in area 1. The estimates for each horizon are separated by two columns (1) without the season fixed effects  $\zeta_s$  and (2) with the season fixed effects. Standard errors are clustered on the season day to account for intra-day correlations across seasons (seasonality). To allow a comparison between the areas only one-day passes are used. The weather and forecast variables are all spatially matched to the same coordinate within the area (lowest lying lift). Easter ( $e$ ) is a dummy indicating the four Easter holidays (Good Friday to Easter Monday). All models are estimated using season day fixed effects to account for the seasonality.

**Table 5: Log demand and its association to weather and forecast variables in area 2**

variables	Forecast d-2		Forecast d-1		Forecast d-0		Weather	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
fsnow	0.0052 (0.0041)	0.0054 (0.0039)	0.0076 (0.0048)	0.0077 (0.0047)	0.010* (0.0048)	0.011* (0.0045)	0.0017 (0.0032)	0.0013 (0.0030)
prec	-0.043 (0.034)	-0.065 (0.034)	-0.044 (0.032)	-0.069* (0.033)	-0.052 (0.038)	-0.068 (0.038)	-0.17*** (0.030)	-0.15*** (0.031)
prec <sup>2</sup>	-0.0034 (0.0035)	-0.0014 (0.0036)	-0.0014 (0.0033)	0.00062 (0.0034)	0.0016 (0.0041)	0.0028 (0.0041)	0.0090** (0.0032)	0.0079* (0.0032)
prec <sup>3</sup>	0.00019* (0.000092)	0.00014 (0.000094)	0.000098 (0.000077)	0.000054 (0.000082)	-0.000027 (0.000100)	-0.000049 (0.000099)	-0.00013 (0.000088)	-0.00012 (0.000089)
sun	0.0091*** (0.0012)	0.0078*** (0.0012)	0.010*** (0.0011)	0.0087*** (0.0011)	0.0096*** (0.0014)	0.0086*** (0.0013)	0.0058*** (0.00077)	0.0066*** (0.00071)
wct	-0.039*** (0.0091)	-0.027** (0.0085)	-0.032*** (0.0082)	-0.023** (0.0080)	-0.034*** (0.0093)	-0.024** (0.0084)	-0.026** (0.0087)	-0.020* (0.0081)
wct <sup>2</sup>	-0.0023*** (0.00051)	-0.0021*** (0.00047)	-0.0019*** (0.00048)	-0.0017*** (0.00046)	-0.0022*** (0.00051)	-0.0020*** (0.00047)	-0.0011* (0.00055)	-0.0013* (0.00050)
easter	1.46*** (0.21)	1.38*** (0.20)	1.42*** (0.19)	1.33*** (0.19)	1.44*** (0.21)	1.35*** (0.20)	1.39*** (0.21)	1.30*** (0.20)
season FE	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>
season day FE	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
<i>N</i>	1099	1099	1099	1099	1099	1099	1099	1099
<i>R</i> <sup>2</sup>	0.677	0.707	0.675	0.707	0.662	0.699	0.658	0.699

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Table Notes:** The table depicts LSDV estimates of model (10) for three forecast horizons and the weather in area 2. The estimates for each horizon are separated by two columns (1) without the season fixed effects  $\zeta_s$  and (2) with the season fixed effects. Standard errors are clustered on the season day to account for intra-day correlations across seasons (seasonality). To allow a comparison between the areas only one-day passes are used. The weather and forecast variables are all spatially matched to the same coordinate within the area (lowest lying lift). Easter ( $e$ ) is a dummy indicating the four Easter holidays (Good Friday to Easter Monday). All models are estimated using season day fixed effects to account for the seasonality.

**Table 6: Log demand and its association to weather and forecast variables in area 3**

variables	Forecast d-2		Forecast d-1		Forecast d-0		Weather	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
fsnow	0.0022 (0.0039)	0.00094 (0.0037)	0.0060 (0.0047)	0.0050 (0.0043)	0.0055 (0.0052)	0.0032 (0.0051)	0.0012 (0.0044)	0.0019 (0.0040)
prec	-0.089** (0.028)	-0.083** (0.026)	-0.12*** (0.029)	-0.12*** (0.029)	-0.077** (0.029)	-0.069* (0.029)	-0.20*** (0.034)	-0.19*** (0.034)
prec <sup>2</sup>	0.0017 (0.0022)	0.0015 (0.0020)	0.0033 (0.0028)	0.0035 (0.0028)	0.00059 (0.0033)	-0.00010 (0.0033)	0.0088 (0.0050)	0.0080 (0.0051)
prec <sup>3</sup>	-0.000020 (0.000035)	-0.000020 (0.000033)	-0.000043 (0.000059)	-0.000052 (0.000060)	0.0000067 (0.000086)	0.000018 (0.000085)	-0.00011 (0.00017)	-0.00010 (0.00017)
sun	0.010*** (0.0011)	0.011*** (0.00099)	0.0091*** (0.0011)	0.0098*** (0.0010)	0.010*** (0.0011)	0.011*** (0.0010)	0.0093*** (0.00059)	0.010*** (0.00060)
wct	-0.015* (0.0060)	-0.011 (0.0061)	-0.017** (0.0055)	-0.011 (0.0057)	-0.016** (0.0057)	-0.0089 (0.0059)	-0.010* (0.0044)	-0.0052 (0.0046)
wct <sup>2</sup>	-0.0022*** (0.00040)	-0.0017*** (0.00042)	-0.0022*** (0.00035)	-0.0016*** (0.00039)	-0.0021*** (0.00035)	-0.0014*** (0.00037)	-0.0012** (0.00038)	-0.00058 (0.00039)
1.easter	0.99*** (0.24)	0.92*** (0.24)	0.81** (0.24)	0.73** (0.24)	0.90*** (0.26)	0.83*** (0.24)	0.93*** (0.24)	0.86*** (0.23)
season FE	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>
season day FE	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
<i>N</i>	1154	1154	1154	1154	1154	1154	1212	1212
<i>R</i> <sup>2</sup>	0.778	0.796	0.773	0.791	0.754	0.773	0.784	0.804

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Table Notes:** The table depicts LSDV estimates of model (10) for three forecast horizons and the weather in area 3. The estimates for each horizon are separated by two columns (1) without the season fixed effects  $\varsigma_s$  and (2) with the season fixed effects. Standard errors are clustered on the season day to account for intra-day correlations across seasons (seasonality). To allow a comparison between the areas only one-day passes are used. The weather and forecast variables are all spatially matched to the same coordinate within the area (lowest lying lift). Easter ( $e$ ) is a dummy indicating the four Easter holidays (Good Friday to Easter Monday). All models are estimated using season day fixed effects to account for the seasonality.

### 8.3 The bias-variance trade-off of the LSDV and WT estimator

Applying the least-squares dummy variable (LSDV) and the within (WT) estimator to the baseline model (10) does not necessarily imply equal prediction accuracy. To see this, consider the least-squares dummy variable (LSDV) and the within (WT) estimator. The former is estimating (??) by

$$y_{ds} = w'_{ds}\beta + \alpha_1 d_1 + \dots + \alpha_D d_D + e_{ds}\gamma + \varepsilon_{ds}. \quad (24)$$

In this formulation the fixed effect  $\alpha_d$  is represented by 155 dummy variables for each day in a season. It fits a line to the regressors after adding or subtracting for each season day a constant (the "fixed" effect). The  $\hat{\alpha}$ s estimate this constant jump in demand compared to the base category and capture the between season day variation. Thus, the  $\hat{\beta}$ s estimate the within season day variation of weather effects. Using this as a prediction estimator has the caveat that predicting the between variation from at most 10 days across seasons produces a very tight fit and, therefore, leads to a relatively high prediction variance.

The WT estimator, on the other hand, estimates

$$\widetilde{y}_{ds} = \widetilde{w}'_{ds}\beta + \widetilde{e}_{ds}\gamma + \varepsilon_{ds}, \quad (25)$$

where  $\widetilde{w}_{ds} = w_{ds} - \bar{w}_d$ ,  $\widetilde{y}_{ds} = y_{ds} - \bar{y}_d$  and  $\widetilde{e}_{ds} = e_{ds} - \bar{e}_d$ . The upper bar over the variable indicates the mean of that variable across seasons within season days. The WT estimator demeans all regressors by season days and this results in exactly the same  $\hat{\beta}$  as in the LSDV estimator.<sup>24</sup> In contrast to the LSDV estimator, the demeaning purges the between season day variation before estimating (25). Predicting demand using the WT estimator requires therefore only the demeaned weather variables as regressors compared to the additional requirement of 155 dummy variables by the LSDV. The higher degrees of freedom of the WT estimator produce mechanically a worse fit. This results in a much lower variance at the expense of some bias in the mean squared error (MSE) of the prediction (see e.g. Hastie et al. (2008)).

### 8.4 Risk for different forecast horizons

The shorter the forecast horizon the more accurate it becomes and, by that, the risk of skiing in unfavoured weather conditions is reduced. In Figure 13 the root-mean-squared errors (RMSE) are depicted for the weather index (first row) and the three key variables

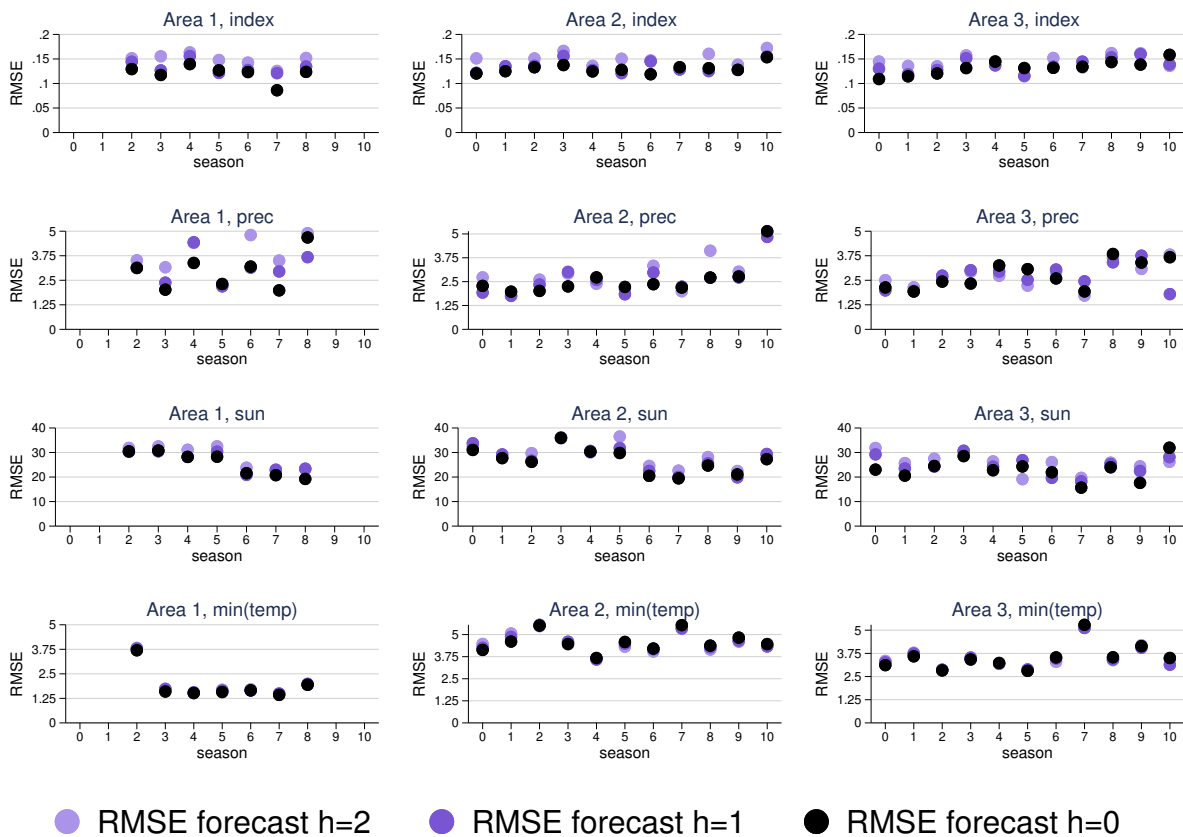
---

<sup>24</sup>By the FWL-theorem, see e.g. Lovell (2008).



in all three areas over all ten seasons. From these graphs we learn three things: (i) In almost all depicted groups the black dots - indicating the shortest forecast horizon - display the lowest RMSE supporting the idea that 0-day forecasts are more accurate than 1- and 2-day forecasts. (ii) The differences between the three forecast horizons are surprisingly low in accuracy - all points are in close proximity and suggest that the weather-related risk reduction between these forecast horizons is very small if not null. (iii) Prediction errors do not follow any systematic trend over the observed periods. The prediction in area 1 became slightly better, whereas in area 2 and 3 it is slightly worse in season 10 than in season 0. But, more strikingly, no clear pattern is visible.

**Figure 13: RMSE of forecast horizons split by season, areas and weather variables**



**Figure Notes:** On the horizontal axes are the seasons depicted and the vertical axes indicate the root-mean-squared errors (RMSE) of different weather variables. The RMSE are depicted for all three forecast horizons  $h = 1, 2$  or  $3$ . In the top row we find the forecast indices, in rows 2 to 4 the three key-weather variables identified in section 3.5.

## 8.5 Derivation of ATT

Suppose that absent any inaccuracies in the forecast outcome is determined by

$$E(y_{ds}^0 | d, s, w_{ds}) = \alpha_d + \beta w_{ds} \quad (26)$$

with  $w_{ds}$  being the weather index,  $\alpha_d$  being the season day fixed effect that is common across seasons but fixed across season days. To evaluate the effect of a wrong forecast on skier outcome, I first define two dummy variables that indicate whether a forecast was optimistic, pessimistic or correct on a given day, denoted by  $D_{ds}^o = \mathbb{1}[(f_{ds}^{d-h} - w_{ds}) > 0.1]$  and  $D_{ds}^p = \mathbb{1}[(f_{ds}^{d-h} - w_{ds}) < -0.1]$ . Now, for those days with an incorrect forecast outcome is determined by

$$E(y_{ds}^1 | d, s, w_{ds}) = \alpha_d + \beta w_{d,s} + \delta^o D_{ds}^o + \delta^p D_{ds}^p$$

which separates into optimistic and pessimistic days

$$\begin{aligned} E(y_{ds}^{1,o} | d, s, w_{ds}, D_{ds}^o = 1) &= \alpha_d + \beta w_{ds} + \delta^o \\ E(y_{ds}^{1,p} | d, s, w_{ds}, D_{ds}^p = 1) &= \alpha_d + \beta w_{ds} + \delta^p \end{aligned}$$

Then, under the conditional independence assumption, the ATTs are given by

$$\begin{aligned} \delta^o &= E(y_{ds}^{1,o} - y_{ds}^0 | d, s, w_{ds}) \\ \delta^p &= E(y_{ds}^{1,p} - y_{ds}^0 | d, s, w_{ds}). \end{aligned}$$

Note that the assignment of the treatments is random as long as the forecaster does not systematically render the forecast either optimistic or pessimistic. In a classical policy evaluation study, the ATT are recovered using a two-way fixed effect model. As our specification is not dependent on the direction of time (simply comparing treated days to untreated is not corrupted by the fact that one day follows the other) and seasonal effects do not alter any coefficients in the baseline specification in section ??, it is sufficient to include individual fixed effects alone (in our case the season day fixed effects  $\alpha_d$ ).

## 8.6 Estimation of ATT

The empirical model to test the above hypotheses is derived using a potential outcome framework (see e.g. Angrist and Kruger) where I compare skiing demand on days with optimistic/pessimistic forecasts to a potential outcome with accurate forecasts. Thus, I

assign days with an optimistic/pessimistic forecast as being treated and days with accurate forecasts as controls. As the forecaster does not systematically publish wrong predictions, the treatments are considered to be random after controlling for seasonality and the weather. The model to be estimated is then

$$\log(y_{ds}) = w_{ds}\beta + \delta^o D_{ds}^o + \delta^p D_{ds}^p + \alpha_d + e_{ds}\gamma + \varepsilon_{ds}, \quad (27)$$

with  $w_{ds}$  being the weather index,  $D_{ds}^o = \mathbb{1}[(f_{ds}^{d-h} - w_{ds}) > 0.1]$  is a dummy indicating optimistic forecast days,  $D_{ds}^p = \mathbb{1}[(f_{ds}^{d-h} - w_{ds}) < -0.1]$  is a dummy indicating pessimistic forecast days,  $\alpha_d$  being the season day fixed effect that is common across seasons but fixed across season days,  $e$  is a dummy indicating Easter holidays and  $\varepsilon$  is the error term. This specification allows that the coefficients  $\delta^o$  and  $\delta^p$  recover average treatment effects on the treated (ATT) days. A short derivation of this result is in Appendix 8.5. Controlling for the actual weather is crucial here because I want to recover the effect of a forecast above/below the variation that is induced by the weather alone.

I estimate model 27 by a least-squares dummy variable estimator (LSDV). The results are presented in Table 7. The estimates of pessimistic forecasts are statistically significant across all areas and most forecast horizons. A pessimistic 0-day forecast is estimated to decrease demand by 39% relative to a day with an accurate forecast in area 1. The same reductions in area 2 and 3 are estimated to be around 25% and 17%, respectively. Optimistic forecasts are only statistically distinguishable from zero in area 2 where a 0-day optimistic forecast increases demand by 26% relative to a comparable day with an accurate forecast. Notice that the effect sizes for the weather index are much larger than for the forecasts. In area 1 I estimate that roughly one in seven skiers decide early as

**Table 7: Effect of optimistic and pessimistic forecast indicators on log demand**

	area 1			area 2			area 3		
	0-day forecast	1-day forecast	2-day forecast	0-day forecast	1-day forecast	2-day forecast	0-day forecast	1-day forecast	2-day forecast
optimistic	0.025 (0.052)	0.10 (0.054)	0.069 (0.053)	0.28*** (0.057)	0.26*** (0.061)	0.20** (0.062)	0.093 (0.050)	0.045 (0.048)	0.083 (0.043)
pessimistic	-0.32*** (0.063)	-0.15* (0.064)	-0.17** (0.057)	-0.24*** (0.063)	-0.20** (0.066)	-0.20*** (0.057)	-0.18** (0.057)	-0.20** (0.062)	-0.16** (0.051)
weather_index	0.023*** (0.0013)	0.025*** (0.0013)	0.025*** (0.0014)	0.020*** (0.0015)	0.020*** (0.0015)	0.020*** (0.0015)	0.029*** (0.0011)	0.029*** (0.0011)	0.029*** (0.0011)
easter	1.02*** (0.19)	0.98*** (0.19)	0.97*** (0.20)	1.46*** (0.19)	1.40*** (0.19)	1.44*** (0.20)	0.89*** (0.23)	0.86*** (0.23)	0.86*** (0.24)
season day FE	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
<i>N</i>	913	913	913	1099	1099	1099	1212	1212	1212
<i>R</i> <sup>2</sup>	0.788	0.783	0.782	0.670	0.669	0.666	0.788	0.788	0.787

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Table Notes:** The table depicts LSDV estimates of model (27) for three areas and three forecast horizons. To allow a comparison between the areas only one-day passes are used. The same tables for area 2 and area 3 demand in all passes are in Appendix XX (in area 1 only one-day passes are used). The variables optimistic and pessimistic are dummies  $D_{ds}^o = \mathbf{1}[(f_{ds}^{d-h} - w_{ds}) > 0.1]$  indicating optimistic forecast days and  $D_{ds}^p = \mathbf{1}[(f_{ds}^{d-h} - w_{ds}) < -0.1]$  indicating pessimistic forecast days. The weather ( $w_{ds}$ ) and forecast  $f_{ds}^{d-h}$  indices are continuous variables scaled between 0 and 1. Easter is a dummy indicating the four Easter holidays (Good Friday to Easter Monday). All models are estimated using season day fixed effects to account for the seasonality.

**Table 8: Effect of optimistic and pessimistic forecast indicators on log demand**

	area 1			area 2			area 3		
	0-day forecast	1-day forecast	2-day forecast	0-day forecast	1-day forecast	2-day forecast	0-day forecast	1-day forecast	2-day forecast
strongly_optimistic	0.036 (0.13)	0.079 (0.12)	0.044 (0.11)	0.22 (0.12)	0.30* (0.12)	0.20 (0.10)	-0.071 (0.076)	-0.024 (0.078)	-0.017 (0.071)
slightly_optimistic	0.035 (0.060)	0.13* (0.058)	0.060 (0.060)	0.23*** (0.065)	0.22** (0.071)	0.19** (0.068)	0.13 (0.077)	0.12 (0.072)	0.086 (0.079)
slightly_pessimistic	-0.23** (0.077)	-0.072 (0.083)	-0.14 (0.085)	-0.22*** (0.053)	-0.23*** (0.054)	-0.21*** (0.060)	-0.15** (0.045)	-0.13** (0.046)	-0.13* (0.054)
strongly_pessimistic	-0.50*** (0.11)	-0.27* (0.12)	-0.18 (0.094)	-0.32** (0.11)	-0.21 (0.12)	-0.27** (0.091)	-0.20* (0.079)	-0.26*** (0.076)	-0.18** (0.068)
weather_index	2.32*** (0.13)	2.49*** (0.13)	2.50*** (0.14)	2.01*** (0.14)	2.08*** (0.15)	2.06*** (0.15)	2.94*** (0.11)	2.91*** (0.11)	2.95*** (0.11)
easter	1.01*** (0.19)	0.97*** (0.19)	0.98*** (0.20)	1.49*** (0.19)	1.46*** (0.18)	1.44*** (0.20)	0.91*** (0.24)	0.89*** (0.23)	0.88*** (0.24)
day FE	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
<i>N</i>	913	913	913	1099	1099	1099	1212	1212	1212
<i>R</i> <sup>2</sup>	0.790	0.784	0.782	0.671	0.669	0.667	0.789	0.790	0.788

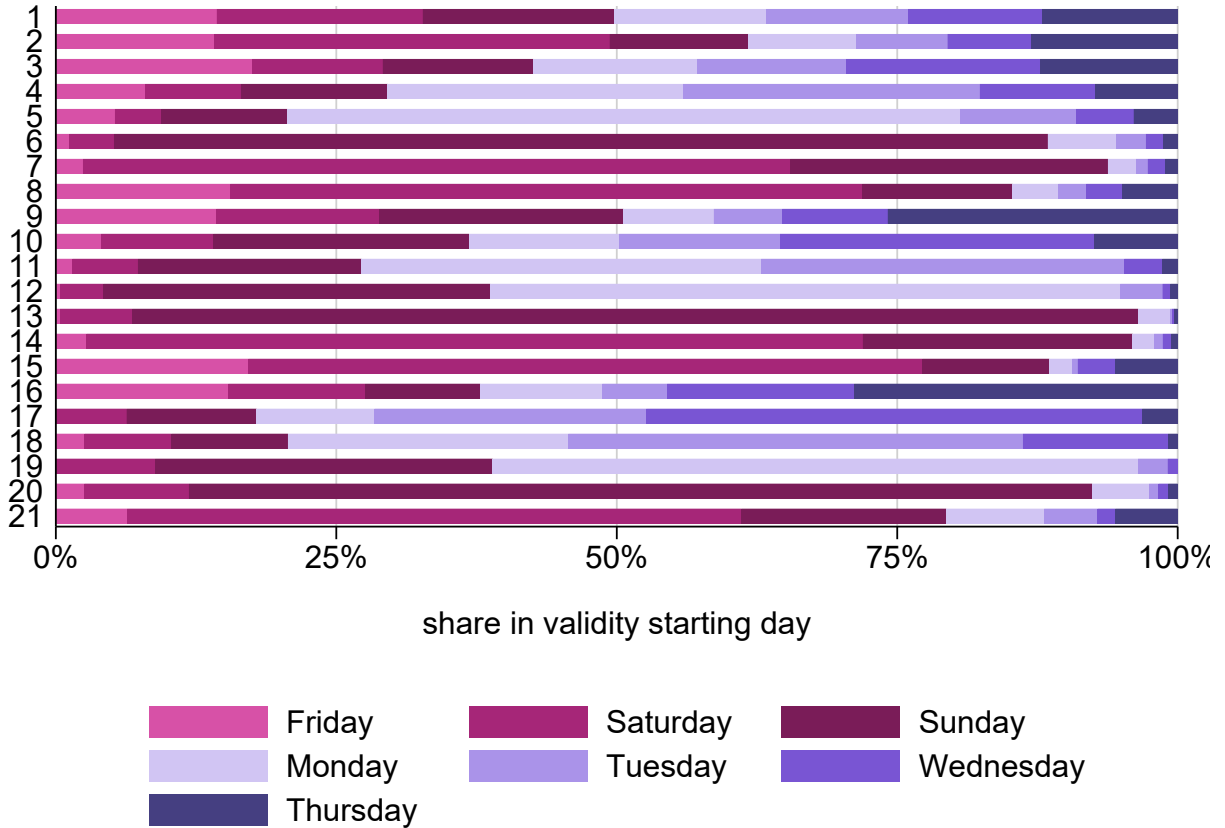
Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Table Notes:** The table depicts LSDV estimates of an extension to model (27) for three areas and three forecast horizons. To allow a comparison between the areas only one-day passes are used. The same tables for area 2 and area 3 demand in all passes are in Appendix XX (in area 1 only one-day passes are used). The variables slightly optimistic and slightly pessimistic are dummies  $D_{ds}^o = \mathbb{1}[0.1 < (f_{ds}^{d-h} - w_{ds}) \leq 0.2]$  indicating slightly optimistic forecast days and  $D_{ds}^p = \mathbb{1}[-0.1 > (f_{ds}^{d-h} - w_{ds}) \geq -0.2]$  indicating slightly pessimistic forecast days. The variables strongly optimistic and strongly pessimistic are dummies  $D_{ds}^o = \mathbb{1}[(f_{ds}^{d-h} - w_{ds}) > 0.2]$  indicating strongly optimistic forecast days and  $D_{ds}^p = \mathbb{1}[(f_{ds}^{d-h} - w_{ds}) < -0.2]$  indicating strongly pessimistic forecast days. The weather ( $w_{ds}$ ) and forecast  $f_{ds}^{d-h}$  indices are continuous variables scaled between 0 and 1. Easter is a dummy indicating the four Easter holidays (Good Friday to Easter Monday). All models are estimated using season day fixed effects to account for the seasonality.

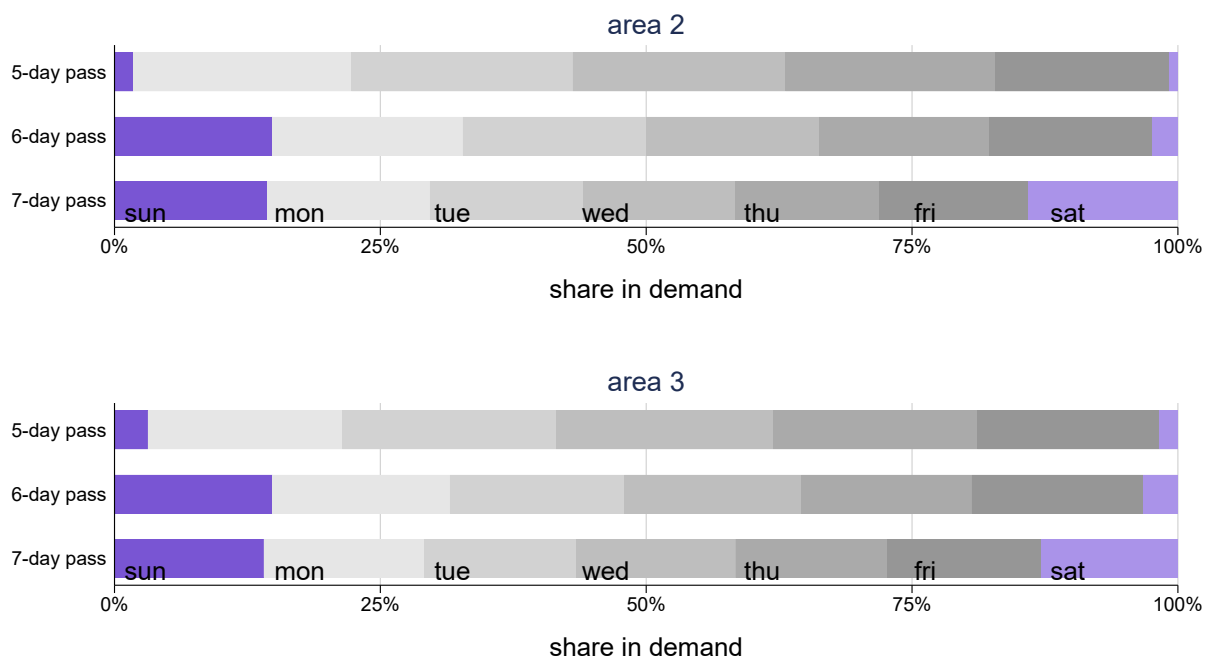
## 8.7 Starting day of validity spells

Figure 14: Shares in starting days of different validity spells in area 1



**Figure Notes:** On the horizontal axis is the share of each starting day in validity depicted. The vertical axis indicates different validity spells from one-day passes up to 21-day passes. Note that five- to seven-day passes often start at a day to cover the weekdays Monday to Friday as most owners of such passes spend their vacation from weekend to weekend.

**Figure 15: Shares in demand for different validity spells in area 2 and 3**



**Figure Notes:** On the horizontal axis is the share of demand for each weekday depicted. The vertical axis indicates different validity spells from five-day passes up to 7-day passes. Note that demand of five- and six-day passes is often high on the days Sunday/Monday to Friday as most owners of such passes spend their vacation from weekend to weekend and depart on Saturday.