



University of Pennsylvania
ScholarlyCommons

Marketing Papers

Wharton Faculty Research

4-2007

Clouds Make Nerds Look Good: Field Evidence of the Impact of Incidental Factors on Decision Making

Uri Simonsohn
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/marketing_papers

 Part of the [Applied Behavior Analysis Commons](#), [Cognition and Perception Commons](#), [Cognitive Psychology Commons](#), [Educational Psychology Commons](#), [Higher Education Commons](#), and the [Marketing Commons](#)

Recommended Citation

Simonsohn, U. (2007). Clouds Make Nerds Look Good: Field Evidence of the Impact of Incidental Factors on Decision Making. *Journal of Behavioral Decision Making*, 20 (2), 143-152. <http://dx.doi.org/10.1002/bdm.545>

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/marketing_papers/288
For more information, please contact repository@pobox.upenn.edu.

Clouds Make Nerds Look Good: Field Evidence of the Impact of Incidental Factors on Decision Making

Abstract

Abundant experimental research has documented that incidental primes and emotions are capable of influencing people's judgments and choices. This paper examines whether the influence of such incidental factors is large enough to be observable in the field, by analyzing 682 actual university admission decisions. As predicted, applicants' academic attributes are weighted more heavily on cloudier days and non-academic attributes on sunnier days. The documented effects are of both statistical and practical significance: changes in cloud cover can increase a candidate's predicted probability of admission by an average of up to 11.9%. These results also shed light on the causes behind the long demonstrated unreliability of experts making repeated judgments from the same data.

Keywords

naturalistic decision making, incidental emotions, priming, college admissions, weather, feature priming, bootstrapped experts, field data

Disciplines

Applied Behavior Analysis | Business | Cognition and Perception | Cognitive Psychology | Educational Psychology | Higher Education | Marketing

Running Head: Clouds Make Nerds Look Good

**Clouds Make Nerds Look Good: Field Evidence of the Impact of
Incidental Factors on Decision Making**

Uri Simonsohn
The Wharton School – University of Pennsylvania

Correspondence Address:

Uri Simonsohn
500 Huntsman Hall
3730 Walnut Street
Philadelphia, PA19104

email: uws@wharton.upenn.edu

Abstract

Abundant experimental research has documented that incidental primes and emotions are capable of influencing people's judgments and choices. This paper examines whether the influence of such incidental factors is large enough to be observable in the field, by analyzing 682 actual university admission decisions. As predicted, applicants' academic attributes are weighted more heavily on cloudier days, and non-academic attributes on sunnier days. The documented effects are of both statistical and practical significance: changes in cloudcover can increase a candidate's predicted probability of admission by an average of up to 11.9%. These results also shed light on the causes behind the long demonstrated unreliability of experts making repeated judgments from the same data.

INTRODUCTION

A growing body of research has studied the impact of incidental and irrelevant factors on judgment and decision making. Within this literature, two somewhat independent streams have studied the role of incidental cognitive primes and incidental emotions. In terms of the former, people's behavior has been shown to be influenced by the presentation of primes in a manner that's consistent with them. In a well known study, for example, subjects primed with words associated with the elderly approached the elevator outside the lab where the study took place at a slower pace than a control group (Bargh, Chen, & Burrows, 1996). For a review of this literature see (Ferguson & Bargh, 2004).

A related line of work has documented that priming people's identity influences their choices. For example, LeBoeuf & Shafir (2003) find that subjects whose 'academic self' was primed were more likely to choose an 'academic' magazine (e.g. The Economist) and Mandel (2003) that subjects primed with their interdependent self (i.e. their reliance on others) become more risk seeking with financial decisions and more risk averse with social ones. Another line of work has documented a phenomenon labeled "feature priming", which consists of primed attributes receiving greater weight in multiattribute decisions or judgments (Mandel & Johnson, 2002; Yi, 1990).

A mostly independent and much more voluminous research stream has documented the impact of (incidental) emotions on judgment and choice. For reviews see Schwarz (2000), Forgas (1995) and Loewenstein & Lerner (2002), and/or the special issue of this journal from April of 2006.

Three main mechanisms have been proposed for the influences of emotions (incidental or otherwise) on judgment and choice. First, emotions influence *how* information is processed. Most importantly for the present research, happy moods induce more heuristic and sad moods more analytical information processing, see (Schwarz, 2002) for a review. Second, emotions enhance accessibility of mood-consistent memories, and third, they provide information (that can be misattributed to the wrong cause if the actual one is not salient). This latter mechanism is often referred to as *mood-as-information*.

Summarizing any one of these three lines of research would require an entire paper, but the following examples of each of the mechanisms are illustrative: (i) Bodenhausen (1993) finds that subjects in happy moods are more likely to rely on stereotypes in the formation of judgments, (ii) Bower (1981) that subjects better recalled words learnt under their current mood, and (iii) Schwarz & Clore (1983) that respondents interviewed on sunnier days express higher levels of overall happiness.

Ultimately, however, such influences of incidental factors are of practical importance only to the extent that they have a sizeable influence on how people make decisions in their everyday lives. If people are influenced by incidental factors only when making hypothetical or low stakes decisions in contrived environments artificially created by an experimenter, but not when making (i) real and important decisions, (ii) in their natural environments, (iii) where they have incentives to make correct choices and (iv) where experience has given them an opportunity to learn how to ignore irrelevant factors, normative theories of choice may still be our best tool for explaining behavior outside the lab.

This paper seeks to assess whether the impact of incidental factors is sufficiently large to be observable and relevant in such a setting, and furthermore, to shed light on the size of the effects they generate in everyday decision making. It seeks, in other words, to test the statistical and practical significance of incidental factors in the field.

To this end, this paper assesses the impact of an ever-present, irrelevant and random incidental factor, cloudiness, on an important and repeated decision, made by professionals in their everyday work environment: university admissions. In particular, this paper analyzes the admission recommendations made for 682 undergraduate applications and assesses the impact of cloudcover the day an application happened to be reviewed, on the weight the reviewers placed on the academic and non-academic attributes of the applicants.¹

Cloudcover has often been studied as a natural manipulator of mood. Prior research, for example, has shown that sunshine increases tipping (Rind, 1996; Rind & Strohmetz, 2001), is positively correlated with returns in the stock-market (Hirshleifer & Shumway, 2003), and leads to increases of self reported levels of happiness (Schwarz & Clore, 1983).

Based on these findings, in Simonsohn (2005) I examined the role of cloudcover during college visits of prospective students on their likelihood to enroll in the visited school. Contrary to initial expectations, visitors on *cloudier* days proved significantly more likely to enroll. I hypothesized that this result may be driven by the fact that cloudcover not only influences people's moods, but also acts as a cognitive prime,

¹ Note that the data consist only of *recommendations*. Data on final decisions are not available.

increasing accessibility to mental constructs which tend to be active during cloudy weather.

Since mellow activities like reading or studying are more appealing and common under cloudy weather and recreational and social activities under sunny weather, it was hypothesized that cloudy weather may prime the former and sunny weather the latter. Because of feature priming, in turn, visitors during cloudy days would weight the school's forte more heavily, academics, while visitors on sunny days would pay more attention to its much weaker social life and entertainment opportunities.²

Support for the hypothesis that cloudy and sunny weather are associated with those two different categories of mental constructs was obtained in a follow up experiment where participants were randomly assigned to a cloudy or sunny weather-forecast prime, and then took part in a word-fragment completion task. Subjects primed with a cloudy forecast were better at solving academic related words like *book* or *student* but not neutral words like *carpet* and *girl*.

Based on this hypothesized link between cloudcover and academics vs. non-academic mental constructs, paired with the notion of feature priming, it was predicted that *college admission reviewers would increase the weight placed on the academic attributes of applicants evaluated on cloudier days, and increase it for the non-academic attributes on those evaluated on sunnier ones*. As is discussed in detail below, however, two of the three mechanisms by which emotions influence choice make the same prediction.

² Although the identity of the school cannot be disclosed, a recent college guide's description is telling of its strengths and weaknesses: "Friends, Sleep, Work, choose two".

First, in terms of the influence of mood on processing style, the literature generally shows increased analytic processing under sad moods, with greater focus in detail (Schwarz, 2002). Happy moods, in contrast, foster increased heuristic processing, broader categorizations and the consideration of a wider range of inputs. This mechanism also predicts, therefore, that on cloudier/sad/focused days, reviewers will place more weight on attributes more closely related to the decision (i.e. academic attributes) while on sunny/happy/inclusive days they will increase their attention to non-academic attributes.

In terms of the priming role of emotions, since cloudcover influences mood, we should expect that high levels of cloudiness will increase accessibility of mental constructs typically experienced under sad moods. To the extent that there is an association between a more mellow emotional state and mental constructs related to academics and/or a more happy/aroused mood and social/fun/non-academic ones (a plausible though untested possibility), this mechanism (paired with feature priming) would also predict that reviewers will place additional weight on applicants academic attributes on cloudy days and on their non-academic ones on sunny ones. Cloudcover, then, may prime academics both directly, and indirectly via mood.

The mood-as-information mechanism does not make any obvious predictions in terms of attribute weighting. It would possibly predict that reviewers, after misattributing their sadder moods to candidates evaluated on cloudy days and their happier moods to candidates evaluated on sunny days, would be less likely to admit students on cloudier days. Daily admission rates would hence be predicted to be negatively correlated with

cloudcover (as we shall see, however, this prediction was not supported by the data as cloudcover has no main effect on admission rates).

Documenting an influence of cloudcover on attribute weighting in actual decisions made by experts would not only demonstrate the practical importance of incidental factors research, but also contribute to the literature investigating the unreliability of expert judgment. Abundant research has shown that experts make inconsistent judgments when making repeated analyses of the same data, for a review see (Ashton, 2000). It is typically assumed that such unreliability is caused by unpredictable factors like fatigue, boredom and distraction (Dawes, Faust, & Meehl, 1989). The results from this paper demonstrate that in addition to the random noise provoked by these elements, incidental factors introduce *systematic* biases which even more strongly argue for the employment of systematic information integration (e.g. simple linear models).

METHOD

Data description

The dataset consists of a sample of 682 paper forms used in the admission process by the university that facilitated the data. These forms are used by admissions' personnel to summarize information about the applicants. Each form contains:

- (i) sixteen 1-4 scores summarizing the applicant's attributes. These ratings are categorized into *academic* (e.g. GPA), *social* (e.g. leadership) and *special consideration* (e.g. outstanding athlete) categories.
- (ii) the admission recommendation of each of two reviewers assigned to review the application, and

- (iii) the date when the application was reviewed by each of the two reviewers.

Variables

Admission recommendations.

The sample contains reviews by at least 15 different reviewers.³ Any given application was evaluated by a subset of two of them. Each of the 682 applicants in the data, then, received two separate admission recommendations for a total of 1,364 observations. Reviewers disagreed on 119 of the 682 applications. Admission recommendations were coded as 1 when a reviewer recommended admission and 0 otherwise.

The total number of applications reviewed per day, was not correlated with cloudcover ($r = .062, p = .598$), suggesting that an influence of cloudcover on attribute weighting is not mediated by effort or fatigue. Another concern is a possible systematic difference in the cloudcover experienced by different reviewers. The F-test test from a regression with admission recommendations as the unit of observation, cloudcover as the dependent variable and reviewers' identities as the only predictors failed to reach significance however ($p = .41$), which means that different reviewers worked experiencing the same average levels of cloudcover.

Finally it is worth mentioning that the offices where applications are reviewed all have windows, providing ample opportunity for cloudcover to be perceived by reviewers.

³ The only personal identifier for reviewers is their handwritten initials. Some initials appear very few times in the data, suggesting they may correspond to coding errors rather than to different reviewers. There are 15 sets of initials with a high enough frequency (17 or more) to strongly suggest they indeed correspond to different reviewers.

Cloudcover

Cloudcover data for the city where the university is located was downloaded from the National Oceanic and Atmospheric Administration (NOAA) website. Cloudcover is measured on a discrete scale from 0 to 10, where 0 is clear skies and 10 is complete overcast. The cloudcover dataset was matched to the admissions dataset based on the date when applications were reviewed. All 11 different values of cloudcover were observed in the sample. The average cloudcover in the data was 7.91 with a standard deviation of 2.32.

Considering that some of the analyses will concentrate on differences in cloudcover experienced by two reviewers of the same application, it is worth noting that reviewers receive stacks of several applications at a time which they pass on to other reviewers only once they have all been reviewed. A given application is hence examined by different reviewers on different days.

Around 80% of the applications were reviewed under a different cloudcover. Importantly, there was no significant correlation in the cloudcover experienced by two reviewers of the same application ($r = -.02$, $p = .604$). Each application, therefore, had two independent ‘draws’ of cloudcover.

Applicants’ attributes

For each applicant, the different 1-4 ratings were averaged by category forming an *academic* ($M = 3.05$, $SD = .472$), a *social* ($M = 2.51$, $SD = .478$), and a *special consideration* ($M = 0.97$, $SD = .136$) average. The special consideration average is low

because 64.6% of the applicants did not have any special considerations to speak of, and hence received a score of 0.

Since these 1-4 ratings are written down by the first reviewer, it is important to establish whether they, independently of any possible effects on the weight they receive, are influenced by cloudcover. To do so I estimated regressions with the academic and the non-academic ratings as dependent variables, and with cloudcover on the day of the first review as the key predictor. Cloudcover was not a significant predictor in any of these regressions (p-values of .24, .72 and .27 for academic, social and special rating respectively).⁴

Some of the analyses require a measure of the *relative* academic strength of each applicant. For this purpose a *nerd-index* was constructed; the *nerd-index* consists of an applicant's academic average divided by the social average ($M = 1.255$, $SD = 0.290$).

Analyses

Three closely related analyses were conducted on the data just described. The first compares the profiles of students admitted on cloudy and sunny days. The second bootstraps weights implicitly placed by reviewers on the three ratings (academic, non-academic and special) separately for sunny and cloudy days, and the third estimates such weights through a single regression model which focuses on differences between reviewers of the same applicant.

⁴ For the academic rating I also estimated a regression controlling for GPA from high-school and SAT scores. These two proved, not surprisingly, significant. Cloudcover, however, remained non-significant ($p = .25$).

Profile of students reviewed and admitted

If reviewers increase their weighting of academic attributes on cloudier days and of non-academic ones on sunnier days, then students who are admitted on cloudier days will tend to be relatively stronger academically and those admitted on sunnier days in their non-academic attributes.

To test this prediction the average *nerd-index* (again: applicants' academic rating divided by their social rating) was computed for students admitted on days with cloudcover above and below 5 in the 0-10 cloudcover scale. As predicted, the nerd-index was significantly higher for students admitted on cloudier days ($M = 1.239$, $SE = .011$) than for those admitted on sunnier ones ($M=1.195$, $SE = .019$), $t(235) = 2.05$, $p = .041$.⁵

The average nerd-index of all students reviewed, i.e. of both those admitted and denied admission, on cloudy ($M = 1.251$) and sunny days ($M = 1.264$), in contrast, was not statistically different, $t(1362) = .66$, $p = .501$. This means that the pools of applicants reviewed on days with different cloudcover were statistically identical, yet the subsets of students who were admitted from such pools were significantly different.

Bootstrapped weights

Although reviewers do not explicitly write down the weights they place on different attributes of candidates they examine, it is straightforward to estimate such weights through bootstrapped models, which consists of regressions where the dependent variable is the recommendation of the reviewer and the predictors are the attributes of the

⁵ The results are robust to defining as the threshold for cloudy vs. sunny other plausible numbers like 4,6,7 and 8. t-tests run on the difference of the academic and non-academic ratings separately, however, did not prove significant.

applicants; the parameter estimates of each attribute correspond to the implicit weights reviewers placed on them (Dawes & Corrigan, 1974; Goldberg, 1970).

In what follows, bootstrapped weights are first estimated separately for applications reviewed on sunny and cloudy days, and then they are estimated through a single regression where the effect of cloudcover is captured by an interaction term between cloudcover and each of the three ratings.

Because the dependent variable is bounded, the results presented below were obtained with logistic regressions. Using OLS leads to qualitatively equivalent results.

Separate bootstrapped models for cloudy and sunny days.

In order to intuitively capture the impact of cloudcover on attribute weights, separate regressions were estimated for applications reviewed on days with cloudcover above and below 5. In these two regressions the unit of observation is a reviewer's admission recommendation, the dependent variable is dichotomous, taking the value of 1 if the reviewer recommended admission and 0 otherwise, and the only predictors are the three ratings of the applicant. The results are presented on Table 1.

As predicted, the implicit weight placed on the academic rating was higher for applications reviewed on cloudy days ($B_{\text{academic}}=1.797$) than on sunny days ($B_{\text{academic}}=1.464$), and the non-academic ones were higher on sunny days ($B_{\text{social}}=2.401$ and $B_{\text{special}}=.964$) than on cloudy days ($B_{\text{social}}=1.109$ and $B_{\text{special}}=.277$). I assess the statistical significance of this pattern in the next subsection.

Table 1

A single bootstrapped model with cloudcover as a predictor.

Although conducting separate regressions for cloudy and sunny days has the advantage of providing intuitive and easy to interpret results, such an approach does have its limitations. Most importantly, it does not lend itself to easily controlling for other factors (such as other weather variables and/or time-of-year controls), nor to quantifying an average effect size or taking into account heterogeneity across different applicants. Finally, obtaining significance levels for the differences in parameter estimates across regressions is not straightforward, as it requires taking into account how the standard errors from one regression are correlated with those in the other.

An alternative consists of estimating a single regression where cloudcover is a covariate instead of a variable used to decide on which regression a given observation belongs. An additional advantage of estimating a single regression is that the analysis can concentrate on *differences* between reviewers of the *same* application, effectively eliminating all heterogeneity across different applicants.

The results presented in this subsection were hence obtained from a regression where the unit of observation is an application, the dependent variable is the recommendation of the second reviewer minus that of the first reviewer, and the key predictors are the interactions between the three student ratings (academic, social and special) and the difference in cloudcover experienced by the second reviewer minus that experienced by the first.

Intuitively, the coefficients for these interactions capture how the weighting of attributes differed between reviewers of the same application, as a function of the difference in cloudcover they experienced. Note that the second reviewer does observe

the recommendation of the first, and hence may be influenced by it. This differences approach is therefore a conservative estimate of the true effect of cloudiness on attribute weighting.⁶

Since reviewers' decisions were coded as 1 if recommending admission and 0 otherwise, the difference between reviewers can take only three values (-1,0,1), and hence a logistic regression was estimated. Qualitatively identical results are obtained if the regression is estimated with OLS.

Table 2

The results of the regression just described are presented on column 1 of Table 2. As predicted, and consistent with the results from the previous subsections, the interaction between the difference in cloudcover experience by both reviewers (DIF) and the academic rating is positive, indicating that a reviewer experiencing cloudier weather placed greater weight on the (same) applicants' academic ratings.

The interactions of DIF with social and special ratings, in turn, are negative, indicating that the reviewer experiencing sunnier weather placed greater weight on the (same) applicant's social and special consideration attributes. All three interactions are significant at the 5% level. A Wald test strongly rejects the null that cloudcover has no effect on attribute weighting, i.e. that all three interactions are 0 ($\chi^2(3) = 23.19$, $p < .0001$).

Column 2 in Table 2 adds controls for differences in rain, wind and temperature experienced by the two reviewers of the same application, plus the interactions of these

⁶ In the extreme, suppose the second reviewer always imitates the decision of the first; even if there was a large cloudcover effect on the weights of the first reviewer this regression would estimate no effect.

three variables with the three applicant ratings (for a total of 12 additional control variables). Controlling for additional weather variables strengthened the point estimate of all three interactions of cloudcover with the applicants' attributes. To assess whether other weather variables were also playing a role, joint tests were estimated for each of the additional weather variables (temperature, rain and wind) and the corresponding interactions. All three tests failed to be rejected ($p_{\text{temp.}} = .91$ $p_{\text{rain}} = .78$, $p_{\text{wind}} = .44$).

A possible time-of-year confound

One possible concern with estimating regressions where cloudcover is a predictor is that cloudcover varies systematically through the year (not a lot of it during the summer, plenty in the midst of winter). Without an appropriate control for time-of-year, therefore, cloudcover could be picking up the influence of the timing of application reviews.

Although the plausibility of a time-of-year confound is dramatically reduced for the regressions just presented, since they focus on differences between reviewers of the same application and they control for other weather variables, presumably also correlated with time-of-year, it is interesting to empirically estimate the potential role that seasonality may be playing.

If the correlation between reviewers' decisions and cloudcover is spurious because of a confound with time-of-year, then cloudcover conditions from the same calendar date of *another* year should also predict admission decisions (e.g. cloudcover of September 5th 2004 should be just as useful a predictor for recommendations made on September 5th 2004 as for recommendations made on the same date in 2005).

With this in mind, the regression was also estimated adding as a control cloudcover conditions for the same calendar date of the review but from *other* years. Rather than using cloudcover from a single other year, average cloudcover for every calendar date was computed for the four years preceding the sample. Since averages have less measurement and sampling error, a time-of-year confound story would predict that the dependent variable will be more strongly dependent on the 4-year average than on the same day cloudcover, and that adding the 4-year-aveages should heavily attenuate the effect of cloudcover on the day of the review. The results of this regression which controls for average cloudcover are presented in column 3 of Table 2.

Contrary to the time-of-year story, the point estimates of interest remain practically unchanged. Furthermore, none of the interactions between average cloudcover and student attributes are significant at the 10% level, and the joint test of all of them being zero cannot be rejected ($p=.51$).

Effect Size.

As was discussed in the introduction, one of the advantages of studying incidental factors in the field is that one can assess not only their statistical but also their practical significance. To this end I computed the predicted probabilities of each applicant being admitted if evaluated under the lowest and highest levels of cloudcover. The average absolute difference between these two values was 11.9% which means that, on average, an applicant's predicted probability of being admitted increases by 11.9% if her application is read under optimal vs. worst possible cloudcover. As a benchmark, applicants need to increase their academic rating by 28.5% in order to obtain a similar gain in admission probability (under average cloudcover).

To estimate the average effect size of cloudcover *in the sample*, analogous computations were conducted for the impact of a change in cloudcover of one standard deviation (i.e. to a change in cloudcover of 2.32), instead of one equivalent to the full spectrum of the cloudiness scale. This led to an average difference in predicted admission probability of 2.7%, equivalent to increasing the academic rating of the average applicant by 7.4%

CONCLUSIONS

Abundant experimental research has shown that incidental primes and emotions can influence behavior. Such findings impose a serious challenge to normative theories of decision making which assume people engage in optimal usage of information and have stable and well known preferences. For any departure of normative models to be of practical relevance, however, its consequences must be large enough to be observable in real decisions, where decision makers, unlike subjects in experiment, have experience with the decision, have incentives to make the right decision, and face a naturally occurring incidental factor.

This paper provides evidence consistent with consequences of incidental factors being large enough to matter in such situations. It shows that professional university admission reviewers weight the attributes of applicants differently, depending on how cloudy the day is when they happen to be reviewing them. If cloudcover, an unstable, irrelevant, and unavoidable cue is capable of influencing the weighting of attributes on the part of experts working under everyday conditions, it is hard to imagine a situation where human judgments or choices could be free of such influences.

These results also contribute to the literature that has examined the reliability of expert judgment (Ashton, 2000). Experts have repeatedly been shown to make *differing* judgment when analyzing *the same data* on different occasions; their unreliability often blamed on fatigue, boredom and distraction (Dawes et al., 1989). This paper suggests that experts are also unreliable integrators of information because situational factors influence the relative weight placed on different attributes. This is important because unlike the previously suspected causes, the role of incidental factors introduces predictable bias, which is arguably even more malignant. The results provide further justification for employing simple linear models to make repetitive integration of information. OLS, after all, will not change its regression weights in response to the amount of natural light entering the room where it is run.

Table 1. *Logistic regression bootstrapping weights for reviews made on days with cloudcover above and below 5, on 0 (clear skies) to 10 (complete overcast) scale*

Dependent variable: 1 if reviewer recommended admission, 0 otherwise.

	Review performed on	
	cloudy day	sunny day
Intercept	-8.326** (0.639)	-10.756** (1.543)
Academic rating	1.786** (0.178)	1.456** (0.395)
Social rating	1.123** (0.154)	2.410** (0.386)
Special rating	0.282** (0.140)	0.950** (0.326)
Number of observations	1,103	261

*,** significant at the 10% and 5% level respectively

Standar errors below parameter estimates.

Table 2. Logistic Regression of Difference Between Raters

Dependent Variable: Second rater's admission recommendation minus first's (possible values: -1,0,1)

	(1)	(2)	(3)
	Baseline	Adds other weather variables	Adds 4-year average of cloudcover
Intercept 1	-2.330** (0.136)	-2.386** (0.141)	-2.387*** (0.142)
Intercept 0	2.547** (0.136)	2.547** (0.149)	2.580*** (0.153)
Difference in cloudcover between Raters (DIF) ¹	0.148 (0.216)	0.047 (0.255)	0.064 (0.258)
DIF * Academic rating	0.134** (0.064)	0.183** (0.075)	0.167** (0.076)
DIF * Social Rating	-0.191** (0.063)	-0.209** (0.076)	-0.198** (0.077)
DIF * Special Consideration Rating	-0.068** (0.021)	-0.072** (0.025)	-0.068*** (0.025)
Difference in Rain, Wind and Temperature, plus interactions (df=12)	No	Yes	Yes
Four-year-average of DIF for calendar dates of reviews, plus interactions (df=4)	No	No	Yes
Number of observations	682	682	682

Standard errors below parameter estimates

*,**,** significant at the 10%, 5% and 1% level respectively

¹Cloudcover is measured in a 0 (clear skies) to 10 (complete overcast) scale

References

- Ashton, R. H. (2000). A review and analysis of research on the test-retest reliability of professional judgment. *Journal of Behavioral Decision Making*, 13(3), 277-294.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71(2), 230-244.
- Bodenhausen, G. V. (1993). Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition and stereotyping* (pp. 13-37). San Diego, CA: Academic Press.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36(2), 129-148.
- Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81, 95-106.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical Versus Actuarial Judgment. *Science*, 243(4899), 1668-1674.
- Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Sciences*, 8(1), 33-39.
- Forgas, J. P. (1995). Mood and Judgment - the Affect Infusion Model (Aim). *Psychological Bulletin*, 117(1), 39-66.
- Goldberg, L. R. (1970). Man Versus Model of Man - Rationale, Plus Some Evidence, for a Method of Improving on Clinical Inferences. *Psychological Bulletin*, 73(6), 422-432.
- Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance*, 58(3), 1009-1032.
- Leboeuf, R., & Shafir, E. (2003). *Alternating selves and conflicting choices: Identity salience and preference inconsistency*. Unpublished manuscript.
- Loewenstein, G., & Lerner, J. S. (2002). The role of affect in decision making. In R. Davidson, K. Scherer & H. Goldsmith (Eds.), *Handbook of Affective Science*. New York: Oxford University Press.
- Mandel, N. (2003). Shifting selves and decision making: The effects of self-construal priming on consumer risk-taking. *Journal of Consumer Research*, 30(1), 30-40.
- Mandel, N., & Johnson, E. J. (2002). When Web pages influence choice: Effects of visual primes on experts and novices. *Journal of Consumer Research*, 29(2), 235-245.
- Rind, B. (1996). Effect of beliefs about weather conditions on tipping. *Journal of Applied Social Psychology*, 26(2), 137-147.
- Rind, B., & Strohmets, D. (2001). Effect of beliefs about future weather conditions on restaurant tipping. *Journal of Applied Social Psychology*, 31(10), 2160-2164.
- Schwarz, N. (2000). Emotion, cognition and decision making. *Cognition and Emotion*, 14(4), 433-440.
- Schwarz, N. (2002). Situated Cognition and the Wisdom of Feelings: Cognitive tuning. In B. Feldman, L. & P. Salovey (Eds.), *The Wisdom of Feelings: Psychological Processes in Emotional Intelligence* (pp. 144-166). New York: Guilford.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality & Social Psychology*, 45(3), 513-523.
- Simonsohn, U. (2005). Weather to Go to College (not a typo). *Working paper*.

Yi, Y. J. (1990). The Effects of Contextual Priming in Print Advertisements. *Journal of Consumer Research*, 17(2), 215-222.