

Recreation of Temperature and Decisions: Evidence from 207,000 Court Cases

Viktor Reif

December 31, 2021

Abstract

(note: cursive indicates my comments) 0. *Abstract: summarize the key points of the paper.* I replicate the main results of the paper "Temperature and Decisions: Evidence from 207,000 Court Cases" by Heyes and Saberian (2019) and evaluate the validity of their analysis. The paper's underlying specification to model the effect of outdoor temperature on asylum court case outcomes is a panel model including control variables for weather, pollution and case characteristics as well as several fixed effects for entities, time and space dynamics. I find the same effects as Heyes and Saberian. Yet through a more thorough analysis (sampling, specification and data issues) get results that question the validity of the original findings. Using yearly subsamples of the entire dataset (2000-2004) yields temperature effects that are not significantly different from zero for all but one year.

Keywords: decision-making; temperature; fixed-effects regression; spatial panel data

1 Introduction

introduce the topic and state the aim of your work, stating clearly the research questions and the methodology used, give a brief overview of the results and the limitations of your analysis. a what my paper does, b what underlying paper does (and why relevant?), c what i can confirm reject

This paper examines the robustness of the results in the article "Temperature and Decisions: Evidence from 207,000" by Heyes and Saberian in 2019. Using the same dataset, this paper recreates the main findings. The aim of this paper is to either empirically confirm the results of Heyes and Saberian or to disprove them and illuminate the reasons for that. I analyse the dataset in Python, recreate the most relevant tables from the original paper in Stata and implement some variations to further analyze the validity of the results.

Heyes and Saberian (2019) use a dataset of 207,000 court cases in the U.S. for a holistic regression analysis to evaluate the influence of outdoor temperature on professionally made decisions. The authors use a holistic set of explanatory variables including various fixed effects - over time, across judges and locations, etc - to control for heterogeneity in the regression of court case outcome on temperature. In this analysis they find a significant relation between temperature and likeliness that a case has a positive outcome, meaning that asylum is granted.

I am able to replicate the paper's main finding. In my analysis, estimated coefficients are very similar in values and perfectly equal in direction and significance. Moreover, I find that the underlying analysis has some data issues, which if taken care of, lead to results that do not contradict the main finding. A more critical finding on my side was that taking subsamples of the dataset yields insignificant temperature effects. This finding strongly questions the validity of the original results. *vllt so wie in papern noch aufzählen was ich danach mach*

2 Literature review

introduce the topic and state the aim of your work, stating clearly the research questions and the methodology used, give a brief overview of the results and the limitations of your analysis. a what my paper does, b what underlying paper does (and why relevant?), c what i can confirm reject summarize the current state of knowledge on the topic of your paper (including the latest relevant publications). all post 2019

Heyes and Saberian already give an exhaustive overview in their paper from 2019. Their work is in line with numerous publications showing that temperature - both indoors and outside - does have a significant effect on human decisions and rationality. More recently, in this branch of temperature x decisions literature Gavresi et al. (2021) show that higher outdoor temperature increases risk appetite in (optimist) financial decisions. Chen et al. (2020) find that people perform worse in neurobehavioral cognitive tests when exposed to higher temperature indoors and Hadi and Block (2019) show that extreme heat makes consumers less rational (ie affectual). Even more temperature effects are shown by Stevens et al. (2021) on aggression on social media and by Ryan (2020) on law officials behaviour.

There is also a group of researchers who disprove the link between temperature and decisions, which Heyes and Saberian omit in their paper. Recent contributions in this branch are Stroom et al. (2021), who find no relation between indoor temperature and cognitive rationality, and Liu et al. (2020), who observe no effect of heat on fraudulent behaviour. *(maybe give here some older literature too, as heyес and saberian do not mention it in their literature review)*

Concerning temperature effects on juridical outcomes specifically, Heyes and Saberian are the first to conduct a full empirical analysis. This motivated Evans and Siminski (2021) to do their own empirical analysis about criminal court cases in Australia, which resulted in no significant effect between weather variables and decision making. Also, as direct response to the underlying paper Holger Spamann (2020) recalculate its results within a larger timeframe (1990 - 2019) and finds no significant effects.

3 Data

Describe the main sources of your data, the data cleaning and merging process, include a table(s) of summary statistics and a brief description of these. Note: missing explanation of avg temp units 10°F and bins.

The main dataset is constructed out of several sources. asylumlaw.org contains the law data in the form of the variables case outcome, case type and nationality of applicant structured along the dimensions judge, city and date (*I still have to check here what the unique key is or how i phrase that*). For the environment data, the National Oceanic and Atmospheric Administration yields air temperature, dew point, air pressure, precipitation and wind speed sorted hourly by datetime and location. The variable cloud cover is available at the Northeast Regional Climate Center. The pollution variables quantity of micro particles, carbon monoxide and ozone are delivered by United States Environmental Protection Agency. Some of the environment data is collected hourly and some daily. As the law variables are in a daily format, hourly data is averaged daily form 6AM to 4PM. Each environment observation is at maximum 32 kilometers away from the respective court location. All variables are joined by date (daily) and city in the dataset machet.dta, thus that every row represents one case outcome marked by a respective date and location containing values for case and environment characteristics.

Once matched.dta is created, the (stata) code puts certain temperature variables into promils and creates variables for all relevant dimensions (city, judge, year, month, day), averages of some characteristics across various dimensions, dummies and interactions between variables and/or dummies. The final dataset contains 207207 observations for 666 variables. The issue of missing values is addressed by dropping every row that contains NA for at least one variable used in at least one regression. Note that this reduces the effective number of observations to 169006.

Table 1: Summary Statistics

	Mean	Std. Dev.
res	0.162965	0.369334
tempmean	61.439452	14.859341
heat	57.398058	16.094140
airpressure0	29.661536	0.751446
avgdewpt	49.392714	16.657781
precip0	0.003891	0.034818
windspeed0	6.518397	4.402740
skycover	0.546602	0.280155
ozone	0.021916	0.012003
co	0.930650	0.504708
pm25	14.869682	11.204614

Table 1 shows summary statistics for the most relevant variables. About 16 percent of all cases end in granting the applicant asylum. As noted by Heyes and Saberian, the grant rate differs greatly across judges and location. over the study period in the Los Angeles courthouse there are five judges that granted asylum to fewer than 4 percent while three others granted in over 67 percent. The mean over the entire dataset for daily average temperature is 61.4°F, which is around 14°C.

4 Empirical strategy

describe the empirical method used and its appropriateness in this context, state the main hypotheses to be tested. (checken ob spezifik auch die von table 2 ist... besonders) struktur absätze: 0. main hyp to test: 1. method: pols

with fixed effects: 2. why appropriate: 3. my further analysis (to be continued):

The main hypothesis to be tested is whether outdoor temperature has an impact on professional high-stakes decisions. In a more empiric logic, this hypothesis is tested using a linear probability model for binary response (for a detailed description see Wooldridge (2010)). The probability model allows for each regressor to influence the likelihood that the dependent variable takes the value of 1. A value of 1 means that asylum is granted. The following model tests the main hypothesis,

$$g_{it} = \beta_0 + \beta_1 temp_{it} + W_{it}\beta_2 + P_{it}\beta_3 + X_{it}\beta_4 + \gamma_i + \psi_{ct} + \theta t + \epsilon_{it} \quad (1)$$

where the dimensions i , t and c represent application, date and city respectively. Thus, the regressand g_{it} is the outcome of an application i on the date t having the value 1 if asylum was granted and 0 otherwise. β_j are the j slope parameters of each regressor (or element of respective regressor matrix) and β_0 is used as the intercept. $temp$ is the main regressor of interest, whereas W , P and X are a set of control variables representing weather, pollution (see table 1 for details about both) and case-specific characteristics, such as nationality of applicant. γ , ψ and θ are included to control for judge-specific fixed effects (ie which judge is ruling the case), time fixed effects (weekday and years) and city-by-month effects. ϵ contains unobserved heterogeneity along the dimensions of case and date. This serves to control for time and spatial autocorrelation. The fixed effects model is especially suitable for this analysis for two reasons. Firstly, it yields a handy interpretation for the coefficient of interest, which is in turn also comparable to several other studies, that used the same approach to model decision making or temperature effects. As this is a probability model, the effect on the dependent variable will always be a change in likelihood(%). Moreover, it is easy to determine and interpret significance of all controls and fixed effects. Secondly, this model can include many characteristics fixed in their respective dimensions. Thusly, as done by Heyes and Saberian, the fixed effects model allows for a holistic approach when testing and including numerous fixed effects.

Apart from the normal method, I also propose spatial lag model using the 50 states, and yearly subsamples of the main dataset.

5 Results

present and comment on your results.

Table 2 contains the results of the regression using the default specification. All four regressions are use pooled OLS to estimate the effects of average temperature and its one-day lag as well as lead in different combinations. Also, all specifications control for a set of averaged weather characteristics (skycover, air pressure, wind, precipitation, dewpoint), air pollution (ozone, co, pm) and weekday dummies. In (1), the estimated slope parameter is -1.363. This value means that a 10°F (5.4°C) increase in daily average temperature during a judge decision reduces the probablilty of a positive outcome by 1.363% (ceterus paribus). Considering that the overall average grant rate is 16.3%, 10°F warmer temperature implies a 8.36% (0.01363/0.163) decrease in expected grant rate. This effect is significant at 1% and so are all other effects in in (2) and (3). Analogously to the just interpreted parameter, the lag or lead estimates quantify the effect of the average temperature the day before or after the decision. In (4), all three regressors are used jointly which leads to significant estimates only for lag and lean. This means that in this dataset when controlling for the temperature the day after and before, the daily average temperature when the decision is made has no effect on its outcome. *I am surprised my results differ that much from the original.*

Table 2: Fixed effect estimates: 6 AM - 4 PM average

	(1) base	(2) 1-Daylag	(3) 1-Day lead	(4) all
temp6t410	-1.075*** [0.274]	-1.454*** [0.406]	-1.208*** [0.382]	-1.617*** [0.486]
press6t4	-0.00494 [0.00518]	-0.00500 [0.00518]	-0.00515 [0.00516]	-0.00523 [0.00516]
dew6t4	0.000723*** [0.000213]	0.000765*** [0.000217]	0.000735*** [0.000217]	0.000780*** [0.000222]
prcp6t4	0.0616 [0.0822]	0.0590 [0.0821]	0.0625 [0.0820]	0.0600 [0.0818]
wind6t4	0.000738 [0.000490]	0.000771 [0.000485]	0.000820 [0.000548]	0.000866 [0.000543]
skycover	-0.00292 [0.00501]	-0.00159 [0.00515]	-0.00186 [0.00538]	-0.000343 [0.00551]
ozone	0.493*** [0.160]	0.503*** [0.160]	0.485*** [0.157]	0.494*** [0.157]
co	0.00572 [0.00389]	0.00547 [0.00389]	0.00552 [0.00385]	0.00523 [0.00384]
pm25	-0.00000866 [0.0000987]	-0.0000104 [0.0000986]	-0.0000130 [0.000100]	-0.0000153 [0.0000999]
ltemp6t410		0.361 [0.278]		0.372 [0.277]
letemp6t410			0.139 [0.260]	0.159 [0.260]
<i>N</i>	206924	206924	206924	206924

Standard errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

maybe put this in data chapter: The variable co2 has approx. 50,000 observations missing. See the Figure 1 for time distribution of said characteristic. Heyes and Saberian omit this issue. By including this variable in their analysis they lose a significant part of their sample. Table1A shows the base result without co2 as a control variable. Estimates differ slightly in their values, whereas direction and significance of all effects remain unchanged. Considering that co2 is one of three pollution control variables (eg co2 and pm25 have a 85% correlation), it is likely that excluding one would not change the base result significantly.

Holger Spamann (2020) have already strongly undermined the validity of the results by Heyes and Saberian using a larger scope for their sample selection and not finding significant temperature effects. Table2A lists the first column of the base result (ie base model without lags/leads) for yearly subsamples of the main dataset. As for the year 2001 there is no co2 available in the dataset, it is omitted

for that year. Each yearly result yields negative coefficients for the relation between temperature and grant rate, yet for all years except 2003 this effect is not significant. Hence, applying a smaller scope for the sample selection also puts the original results in question. When combining the findings of Spamann with Table2A, a potential interpretation could be that by chance Heyes and Saberian happened to find significant effects in exactly their sample scope (2000-2004). A short side note here is that the original analysis controls for years with year-dummies. *sind die significant? also not sure about this next sentence...* Assuming those are significant, these fixed effects only control for yearly differences in the dependent variable, but not for differences in the regressors. *mention in discussion oder ka wo yearly subsample vs year fixed effects? mal schauen wie stark da die coefficients sind.*

6 Discussion

reflect on the meaning and policy implications of your results, think of potential limitations to your work and avenues for future research. 1. meaning and implications like in op. 2. potential limits like in op like samplesize, also shortcomings in specification due to var availabilty like op and also methodology (for example latent factors). 3. future research more datasets to see link temp-¿court cases. link more to climate change

The results of this paper are in line with the findings from Heyes and Saberian and so are its implications. The results imply that highly important decisions might not be arbitrary. These shortcomings decrease overall societal welfare and efficiency. As noted in the literature review, other authors have shown that the data selection as done by Heyes and Saberian is flawed. As this analysis uses the same dataset, it suffers from the same obvious limitation. Another limitation might come from the model specification. In that regard there could exist omitted variables that bias the coefficient of interest either due to unavailability of data (variables) or a too parsimonious methodology. Considering for instance each judge's sentiment towards immigration, there is no doubt that omitting this characteristic could cause a bias. Certainly, there exists no dataset for this variable. Also, judge fixed effects are insufficient to control for this surely also time-varying characteristic. The same issue might arise for other unobserved factors, which are all put into the error term and threaten the validity of the model. Mentioning the error term, the analysis lacks any attempts to "mine" (as a general term) the data in question. A feasible idea would be to extend (1) by latent factors as proposed by Bai (2009). This method mines the error term for latent factors (principal components) which, if omitted, would bias the regression. This is just one example to illustrate there are ways to further exploit the dataset to get an estimate for temperature with as little bias as possible.

More research on the link between temperature and decision making is needed to clarify the ongoing debate about it. Especially in the context of court cases more empirical contributions would help, as those are yet scarce. Furthermore, future research can also examine the relation decisions/ human cognitive output and climate change. This is outstandingly relevant as average global temperatures as well as local weather volatility are expected to rise. The relation between decisions and climate change might even be impactful enough to be included in the integrated modelling approaches (IMAGE) of global temperature forecasts. Going back to the basic relation between decisions and current weather, another topic for research would be to mitigate the methodological shortcomings mentioned above. This could be done either through mining techniques to achieve a more adequate model specification, or use other analysis/estimation methods. Examples for the latter would be regression trees or neural networks.

7 Conclusion

7.1 testsubsection

summarize your main work and conclude. 1. Select a paper that uses one of the empirical methods reviewed in class 2. Get the raw data 3. Replicate the data analysis 4. Write a report summarizing your work. 5. Include a literature review section in your report that summarizes the current state of knowledge on your topic.

The main finding of this paper are twofold. On one hand the analysis can confirm the correctness of the empirical findings from Heyes and Saberian. On the other hand a more holistic analysis on my part strongly questions the validity of the underlying paper's findings. Apart from smaller issues, the effects found by Heyes and Saberian do not hold for subsamples (eg yearly intervals) of the main dataset in question. These validity issues found in sampling scopes smaller than the original scope are in line with the finding of other authors. Holger Spamann (2020) uses a larger scope (1990-2019 instead of 2000-2004) and finds no significant temperature effects. Hence, this paper questions the validity of the originally found effects from yet another empirical perspective. With those main findings this paper gravitates towards the stream of publications that negate temperature effects on decision making. Yet this paper also shows several avenues for research to further understand the relation between exogenous variables and human decisions/ cognitive outputs.

Appendix

use it for additional material that might support your analysis + (in the

final version) include a separate paragraph that provides a response to the referee's comments and mentions where, how, why, why not the paper has changed.

A Additional Tables

Table A.1: Fixed effect estimates: 6 AM - 4 PM average

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	nothing	nat	dow	type	judge	cm	city/ym	cym	jm/c/y	date
temp6t410	-1.470*** [0.355]	-0.717*** [0.270]	-0.727*** [0.273]	-0.780*** [0.269]	-0.806*** [0.249]	-1.037*** [0.278]	-0.893*** [0.215]	-0.652** [0.262]	-1.073*** [0.271]	-0.939*** [0.285]
N	206924	206924	206924	206924	206924	206924	206924	206924	206924	206924

Standard errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Fixed effect estimates: 6 AM - 4 PM average (without carbon monoxide as control variable)

	(1) base	(2) 1-Daylag	(3) 1-Day lead	(4) all
temp6t410	-0.877*** [0.260]	-1.237*** [0.365]	-1.147*** [0.352]	-1.532*** [0.454]
press6t4	-0.00535 [0.00470]	-0.00543 [0.00469]	-0.00581 [0.00465]	-0.00592 [0.00464]
dew6t4	0.000638*** [0.000202]	0.000676*** [0.000206]	0.000659*** [0.000206]	0.000699*** [0.000211]
prcp6t4	0.0293 [0.0798]	0.0270 [0.0796]	0.0313 [0.0796]	0.0290 [0.0795]
wind6t4	0.000760* [0.000458]	0.000805* [0.000457]	0.000939* [0.000512]	0.000994* [0.000513]
skycover	-0.00646 [0.00454]	-0.00524 [0.00455]	-0.00426 [0.00475]	-0.00289 [0.00480]
ozone	0.120 [0.132]	0.128 [0.133]	0.105 [0.131]	0.112 [0.132]
pm25	0.0000481 [0.0000965]	0.0000451 [0.0000963]	0.0000363 [0.0000980]	0.0000325 [0.0000979]
ltemp6t410		0.343 [0.261]		0.356 [0.262]
letemp6t410			0.282 [0.248]	0.295 [0.249]
<i>N</i>	250652	250652	250652	250652

Standard errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Yearly results

	(1)	(2)	(3)	(4)
	2000	2002	2003	2004
temp6t410	-0.644 [0.648]	-0.425 [0.509]	-0.928** [0.431]	-0.268 [0.635]
press6t4	-0.00770 [0.0192]	0.0150 [0.0153]	-0.00718 [0.0151]	-0.0140 [0.0173]
dew6t4	-0.0000398 [0.000453]	0.000308 [0.000375]	0.000371 [0.000325]	0.000829* [0.000483]
prcp6t4	-0.00440 [0.157]	0.0848 [0.161]	-0.0245 [0.128]	0.143 [0.203]
wind6t4	-0.000105 [0.000999]	0.000106 [0.000854]	0.00149* [0.000827]	0.000539 [0.000929]
skycover	0.00180 [0.0111]	-0.00775 [0.00905]	-0.00695 [0.00898]	0.00676 [0.00908]
ozone	0.840*** [0.294]	0.566* [0.304]	0.199 [0.258]	-0.0777 [0.317]
co	0.00302 [0.00681]	0.00172 [0.00791]	0.00458 [0.00788]	0.0104 [0.0115]
pm25	0.0000814 [0.000131]	0.0000511 [0.000167]	0.000270 [0.000190]	0.000163 [0.000290]
<i>N</i>	45463	54106	65572	41783

Standard errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Yearly results (without carbon monoxide as control variable)

	(1)	(2)	(3)	(4)	(5)
	2000	2001	2002	2003	2004
temp6t410	-0.644 [0.648]	-0.109 [0.675]	-0.428 [0.509]	-0.943** [0.428]	-0.240 [0.640]
press6t4	-0.00778 [0.0193]	0.0000866 [0.0164]	0.0153 [0.0154]	-0.00638 [0.0153]	-0.0120 [0.0174]
dew6t4	-0.0000237 [0.000453]	0.000538 [0.000587]	0.000317 [0.000371]	0.000395 [0.000328]	0.000844* [0.000486]
prcp6t4	-0.00413 [0.157]	-0.101 [0.209]	0.0854 [0.161]	-0.0228 [0.128]	0.151 [0.203]
wind6t4	-0.000236 [0.000896]	0.00130 [0.00103]	0.0000530 [0.000821]	0.00136* [0.000781]	0.000325 [0.000924]
skycover	0.00131 [0.0110]	-0.0186* [0.00980]	-0.00787 [0.00899]	-0.00709 [0.00895]	0.00606 [0.00898]
ozone	0.841*** [0.293]	-0.459 [0.303]	0.553* [0.293]	0.164 [0.251]	-0.148 [0.309]
pm25	0.0000932 [0.000129]	0.000141 [0.000271]	0.0000607 [0.000162]	0.000310* [0.000178]	0.000215 [0.000283]
<i>N</i>	45463	43728	54106	65572	41783

Standard errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ **Table A.5:** Fixed effect estimates: 6 AM - 4 PM average**Table A.6:** Fixed effect estimates: 6 AM - 4 PM average

References

- Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.
- Chen, Y., Tao, M., and Liu, W. (2020). High temperature impairs cognitive performance during a moderate intensity activity. *Building and Environment*, 186:107372.
- Evans, S. and Siminski, P. (2021). The effect of outside temperature on criminal court sentencing decisions.
- Gavresi, D., Litina, A., and Makridis, C. (2021). Split personalities? behavioral effects of temperature on financial decision-making. *SSRN Electronic Journal*.
- Hadi, R. and Block, L. (2019). Warm hearts and cool heads: Uncomfortable temperature influences reliance on affect in decision-making. *Journal of the Association for Consumer Research*, 4(2):102–114.
- Heyes, A. and Saberian, S. (2019). Temperature and decisions: Evidence from 207,000 court cases. *American Economic Journal: Applied Economics*, 11(2):238–265.
- Holger Spamann (2020). No, judges are not influenced by outdoor temperature (or other weather): Comment.
- Liu, H., Yang, J., and Yamada, Y. (2020). Heat and fraud: evaluating how room temperature influences fraud likelihood. *Cognitive Research: Principles and Implications*, 5(1).
- Ryan, M. E. (2020). The heat: temperature, police behavior and the enforcement of law. *European Journal of Law and Economics*, 49(2):187–203.
- Stevens, H. R., Graham, P. L., Beggs, P. J., and Hanigan, I. C. (2021). In cold weather we bark, but in hot weather we bite: Patterns in social media anger, aggressive behavior, and temperature. *Environment and Behavior*, 53(7):787–805.
- Stroom, M., Kok, N., Strobel, M., and Eichholtz, P. (2021). Turning up the heat: The impact of indoor temperature on cognitive processes and the validity of self-report. *SSRN Electronic Journal*.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.