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

A referee report by Vincent Löwe

Viktor Reif

Name: Vincent Löwe
Matrikel number: 3206198
MSc Economics
Heidelberg University
Supervisor: Jun.Prof. Dr. Anca Balietti

23.01.2022

1 Critical review

Viktor Reif's paper "Recreation of Temperature and Decisions: Evidence from 207,000 Court Cases" is built around the publication of Heyes and Saberian (2019) [2]. Their paper is an addition to the already extensive literature in the fields of psychology and behavioural economics on the effect of environmental factors of a physical nature, such as temperature or humidity, on the decision making process of professionals. To investigate the effects of temperature on the verdict of US judges in asylum application cases, the authors analyse a large data set comprised of a variety of covariates, including a whole array of controls and fixed effects dummy variables, utilising pooled OLS estimation. They ultimately find a statistically significant negative effect of temperature on the decisions made by the judges analysed in the paper.
The vantage point of Reif's paper is his successful replication of the original study, however, he does add some noteworthy facets to it. Most prominently he questions the robustness of the original findings on the basis of similar analyses he conducted on subsets of the original data set, which produce mostly insignificant results. To criticise the paper on the grounds of its sample selection spuriously causing significant results is an interesting point raised by Reif, even though he adopted this critique from a paper published by Holger Spamann [3]. Additionally, he includes a latent factor model into his empirical strategy. This sophisticated instrument, meant to absorb the unobserved heterogeneity that is neither constant over time nor along the other grouping dimensions, certainly is a clever and non-conventional way of dealing with omitted variable bias. However, Reif's reasoning for employing it in this instance is not without its problems.

2 Technical analysis

Reif starts his analysis by replicating the results of Heyes and Saberian [2] with success. To this end he computes a pooled OLS estimation of a linear probability model (LPM) for binary response, which includes a wide range of controls and fixed effects dummy variables, on the data set used by the original authors. The binary dependent variable indicates whether an asylum application was granted

and the variable of interest is the temperature. As touched on in the previous section, Reif fields two major challenges against the results of Heyes and Saberian [2], which I will now discuss in detail:

1. His first challenge is inspired by a publication of Holger Spamann [3], who performs a similar estimation to the one done by Heyes and Saberian [2] on a data set covering a considerably longer time frame, which does not yield any significant results. Spamann takes this as evidence that the sample selection in Heyes and Saberian [2] causes the significant results of the paper. Building on this Reif conducts a series of regressions, in which he performs the same fixed effects estimation on subsets of the original data set pertaining to the different years covered by it. Since the carbon monoxide control variable is completely missing for the year 2001, he repeats the process without the variable controlling for carbon monoxide, so that he can also perform the estimation on the subset for the year 2001. While the estimated coefficients for the temperature covariate were consistently negative, just as in Heyes and Saberian (2019) [2], only those for the year 2003 were significant. On its own this is indeed an interesting result. However, the reduced amount of data might affect the accuracy of the standard errors. The size of the data sets with >45000 data points might still be ample, but the number variables is also considerable. For this reason it is paramount to name the type of standard error computed and defend its accuracy in this setting.

2. Reif also questions the identification strategy of the authors, who believe that the fixed effects included in their estimation strategy are sufficient to combat possible omitted variable bias caused by unobservable heterogeneity. He emphasises the point that there might be unobservable variables that are not time constant and could, thus, still be causing endogeneity. His proposed solution is to account for these possible distortions by estimating a latent factor model. This is achieved by expanding the original estimation model by a matrix of interactive fixed effects v_{mn}. This modified model is subsequently estimated like the baseline specification. The results of this regression are, however, similar in scale and significance to those of the original paper. As an example of a possibly time-variant unobserved omitted variable he

> points to the political sentiment of a judge, which is of course likely to influence decisions on asylum. Now while the discussion about ideological drift of judges is ongoing and undecided [see for instance: 1], even the most adamant proponents of the existence of ideological drift would agree that this process usually takes place over the course of many years or decades. In all likelihood a time frame of 5 years is too narrow for there to be a relevant shift in political convictions. Of course this is just one example and there might be other latent variables one should correct for, but the similar results for his latent factor model could be understood as evidence, that the identification strategy of the original paper actually was sufficient.

One thing Reif could consider to add to his critical analysis of Heyes and Saberian (2019) [2] is a discussion of their choice of a LPM, with its well known deficiencies, such as the possibility of fitted values falling outside of the $[0, 1]$ interval. Instead of adding the latent factor model, which is of questionable use, he might add a logit or probit regression. If he wants to keep controlling for unobservable heterogeneity, he might choose a generalised linear mixed model.

3 Minor Comments

There are some minor orthographic lapses, which should be dealt with. Also it might be conducive to an easier understanding of the paper, if in the tables variables are renamed into something more lucid and different specifications are labelled more clearly.

Total word count: 995

References

[1] Lee Epstein et al. "Ideological drift among supreme court justices: Who, when, and how important". In: *Nw. UL Rev.* 101 (2007), p. 1483.

[2] Anthony Heyes and Soodeh Saberian. "Temperature and Decisions: Evidence from 207,000 Court Cases". In: *American Economic Journal: Applied Economics* 11.2 (Apr. 2019), pp. 238–65. DOI: 10.1257/app.20170223. URL: https://www.aeaweb.org/articles?id=10.1257/app.20170223.

[3] Holger Spamann. "No, Judges Are Not Influenced by Outdoor Temperature (or Other Weather): Comment". In: *Harvard Law School John M. Olin Center Discussion Paper* 1036 (2020).