

June 3, 2024

To: Klaus F. Zimmermann, Editor-in-Chief, Journal of Population Economics

From: Alfonso Flores-Lagunes, Chair of the Ad-hoc Scientific Committee

Re: Report of the Ad-hoc Scientific Committee on Ciacci (2024, Journal of Population Economics)

This document constitutes the report of the Ad-hoc Scientific Committee's findings regarding the paper "Banning the purchase of sex increases cases of rape: evidence from Sweden" by Riccardo Ciacci (hereafter the author), published in the Journal of Population Economics in 2024. This paper was formally challenged by Joop Adema, Olle Folke, and Johanna Rickne (hereafter the critics) through a write-up that was sent to the Editor-in-Chief of the journal, Klaus F. Zimmermann. The Editor-in-Chief appointed one of the editors, Alfonso Flores-Lagunes, to oversee an ad-hoc committee consisting of two experts in the subject matter and himself to evaluate the scientific merits of the situation. The two other members of the committee are Scott Cunningham (Baylor University) who has expertise in the subject matter (economics of sex) and also in causal inference methods; and Sebastian Calonico (Columbia University) who is an expert in regression discontinuity designs, the main method employed in the paper. Alfonso Flores-Lagunes has expertise in causal inference methods and in empirical microeconomic work.¹

The committee received the following information for evaluation: (1) the published article and its online appendix, (2) the write-up from the critics and corresponding replication files, (3) a set of documents that constitute the reply by the author, and which included replication files. **The committee was charged with evaluating the scientific merits of the situation.** This report is divided into three parts: the background of the situation, the findings pertaining to the scientific merits of the case, and some additional thoughts from the committee.

¹ Alfonso Flores-Lagunes was aided by Associate Editor Hugo Jales, who is also an expert on regression discontinuity designs. Hugo Jales independently analyzed the data provided. However, he did not directly participate in the discussions of the committee and its conclusions.

Background

The published paper by the author analyzes the effect on rapes of the introduction of a policy in Sweden that bans the purchase of sex, which took effect in January of 1999. The highlight result in the paper is that the policy increased the number of rapes by between 44% and 62%. The main analysis employs a regression discontinuity design with time as the running variable. The monthly-level data consist of Swedish regions observed from 1997 to 2014.

The fundamental point by the critics of the published article is that the main specification used by the author is incorrect due to the inclusion of year fixed effects and month-of-the-year fixed effects (the latter included to account for seasonality in the number of rapes) that result in multicollinearity with the treatment variable (an indicator for the time when the policy was introduced). The implication of the criticism is that the inclusion of those fixed effects forces the statistical program to arbitrarily drop some fixed effects, which artificially yields the highlight result in the paper.

The reply by the author recognizes that multicollinearity was an issue in his published paper's main specification. He then presents a criterion to choose which fixed effects to drop, and based on this criterion he shows that the results in the original paper still seem to hold under various model specifications.

Findings of the Committee

The committee focused attention on the main method employed—regression discontinuity in time—and the consequences for the conclusions in the paper from the fundamental point brought up by the critics.

First, based on the evidence presented, the committee concluded that the author did **not** engage in “acts of misconduct” as defined in the journal’s policy on “Ethical Responsibilities of Authors” outlined in its [website](#). It is very likely that the issue of multicollinearity was simply not detected by the author, and it was subsequently not detected by the peer reviewers of the paper. Therefore, the committee concludes that this issue constitutes a “fundamental error,” as stated in the previously cited journal’s policy.

Second, the committee agreed with the critics and with the reply of the author that the results in the published paper were affected by an issue of multicollinearity that occurred from the inclusion of multiple sets of fixed effects.²

Third, based on the information provided to the committee, the committee concluded that it is not entirely clear that the results in the published paper would stand after sensibly correcting for the multicollinearity problem. Unfortunately, in his reply, the author did not re-estimate the original models while solving the multicollinearity issue. The estimates from the models presented by the author in his reply—which address several aspects brought up by the critics—are generally of a smaller magnitude and less precise than the highlight result in the published paper. Moreover, results from selected exercises conducted by different members of the committee, including Associate Editor Hugo Jales, indicate a lack of robustness of the highlight result in the published paper. One takeaway, therefore, is that the specific way in which the method is implemented has a substantial impact on the results that are obtained.

Fourth, the committee discussed what would be the most sensible specification to use in this context. The notes here are offered for reference. The point of departure in this discussion is that the idea behind a regression discontinuity in time is to look *locally* at what happens around the “threshold” (the time at which the policy took effect). The method requires that the “running variable” (time) changes smoothly around the threshold. Typically, a (optimal) “bandwidth” is employed to set a neighborhood around the threshold that defines comparable treated and non-treated units (which here are at the level of month-region). Regarding the use of year fixed effects: the presumed smoothness of time and the locality implicit in the method make it unnecessary, in principle, to control for year fixed effects (a non-smooth way to control for time). Instead, it is preferable to use a low-order polynomial to account for the assumed smoothness of time around the threshold. Regarding the use of regional fixed effects: these are not needed given that the regions are perfectly balanced around the threshold. That is, the same regions exist around the threshold, and they are all treated at the same time. Lastly, there is the use of month-of-the-year fixed effects to account for seasonality. Here, the committee agrees that, in general, it can be consequential to account for seasonality if the dependent variable has strong short-term seasonal

² The consequence of the problem of multicollinearity in question is pretty specific given that the running variable is time: the inclusion of the multiple sets of fixed effects results in two month-of-the-year dummy variables (January and December) and two year dummy variables (1997 and 2001) to be dropped by the software.

variation, particularly in months around the threshold (e.g., December to January). This is a case in which the committee thinks that different researchers may account for seasonality differently. At the same time, when looking at these data, the committee finds that seasonality between contiguous months appears not to be strong, implying that accounting for seasonality should not have a strong impact on the resulting estimate. In sum, the committee would consider a “sensible” specification to be one that uses a low-order polynomial in time around the threshold, with no region fixed effects and with the option of accounting for seasonality in a sensible way.

One last word about the “difference-in-discontinuity” specification that the author considers a better model in his reply. The committee found it hard to understand exactly how that model is implemented by the author—it seems to be done in a peculiar way. The committee discussed that a more transparent procedure to implement the “difference-in-discontinuity” specification, which would account for the seasonality in the data, is as follows. First, obtain an effect using the regression discontinuity in time at the real threshold, followed by obtaining a placebo effect using the regression discontinuity in time from 1997 to 1998 (both effects obtained in the same way). Then, subtract the placebo effect—which presumably includes only the seasonality under some assumptions—from the effect using the real threshold to net out the seasonality from the estimated effect using the real threshold.³

In conclusion, the published paper’s results were affected by an arbitrary way of dealing with multicollinearity (performed by the software defaults). Different solutions to this multicollinearity problem yield different answers to the key question that the published paper addresses, and the committee does not think that the author has convincingly established *which* fixed effects ought to be dropped from the model. Also, the highlight result in the paper does not seem robust to how one accounts for the multicollinearity problem, and this estimated effect is likely to be smaller and less precise than was reported in the published paper.

³ The committee recognizes that there are decisions to be made in the implementation of this procedure since there are different number of observations on each side of the threshold and the estimate of the placebo effect cannot include the treated period. As a result, the optimal bandwidth will likely require more observations than are available in each side of the threshold, in each case. The point is that the same implementation decisions be applied to the estimation around the placebo and real thresholds. Still, the committee feels that this procedure can directly address the critics’ point that the published paper only estimates the effect of seasonality (i.e., the “December vs. January” effect).

The Committee's Additional Thoughts

This section is offered as additional thoughts from the committee since they go beyond its charge of evaluating the scientific merits of the situation.

The committee's opinion, informed by the material furnished to it, is that it may be valuable to give the author an opportunity to more systematically assess whether the highlight result of the published paper is robust, particularly in light of the thoughts offered by the committee. This exercise could also include a carefully crafted simulation exercise that can shed light on whether a sensible specification can recover a known average treatment effect. A short summary of the findings from this endeavor could be considered for publication in the journal as a corrigendum and be clearly added to the published version of the article. If this option were considered, the critics could be recognized therein for pointing out the multicollinearity problem in the published article, which is their main contribution.

Last but not least, the committee is saddened by the way the author was treated by the critics in social media and their private email interactions. This is important to state since civility and respect should dominate academic discourse.